

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
16 January 2003 (16.01.2003)

PCT

(10) International Publication Number
WO 03/004689 A2

(51) International Patent Classification⁷: **C12Q**

(US). DRENKARD, Eliana [AR/US]; 50 Massachusetts Avenue, Cambridge, MA 02139 (US).

(21) International Application Number: PCT/US02/21431

(74) Agent: DECAMP, James, D.; Clark & Elbing LLP, 101 Federal Street, Boston, MA 02110 (US).

(22) International Filing Date: 8 July 2002 (08.07.2002)

(25) Filing Language: English

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(26) Publication Language: English

(30) Priority Data:
60/303,286 6 July 2001 (06.07.2001) US
60/373,233 16 April 2002 (16.04.2002) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 60/303,286 (CIP)
Filed on 6 July 2001 (06.07.2001)
US 60/373,233 (CIP)
Filed on 16 April 2002 (16.04.2002)

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): THE GENERAL HOSPITAL CORPORATION [US/US]; 55 Fruit Street, Boston, MA 02114 (US).

Published:

— without international search report and to be republished upon receipt of that report

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): AUSUBEL, Frederick, M. [US/US]; 271 Lake Avenue, Newton, MA 02161

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: REGULATORS OF BIOFILM FORMATION AND USES THEREOF

(57) Abstract: This invention relates to nucleic acid and amino acid sequences of genes regulating bacterial biofilm formation and to the use of these sequences as targets in the diagnosis, treatment, and prevention of bacterial infection and pathogenesis. In addition, the invention relates to screening methods for identifying modulators of bacterial biofilm formation and the development of antibacterial treatments.

WO 03/004689 A2

REGULATORS OF BIOFILM FORMATION
AND USES THEREOF

5

Background of the Invention

This application claims benefit of U.S. provisional applications 60/303,286 and 60/373,233, filed July 6, 2001 and April 16, 2002, respectively. The disclosures of
10 which are hereby incorporated by reference.

This invention relates to nucleic acid and amino acid sequences of genes regulating bacterial biofilm formation and to the use of these sequences as targets in the diagnosis, treatment, and prevention of bacterial infection and pathogenesis. In addition, the invention relates to screening methods for identifying modulators of bacterial biofilm
15 formation and the development of antibacterial treatments.

Bacteria possess the ability to form aggregated, organized, colonial communities called biofilms. Distinct from their free-living planktonic counterparts, bacterial cells that form biofilms secrete an exopolysaccharide slime that surrounds and protects the bacterial colony. By adhering to each other and to surfaces or interfaces, these matrix-
20 enclosed bacterial populations can cause any number of problems. By attaching to a variety of surfaces such as contact lenses, water pipes, hip replacements and food packaging, they can cause irritation, disease, immune rejection, and food poisoning.

In addition to attaching to abiotic surfaces, many biofilm-forming bacteria colonize living tissue where they cause serious infection. For example, *Pseudomonas*
25 *aeruginosa* colonizes the lungs of cystic fibrosis (CF) patients as a biofilm. Chronic colonization of the airways by this bacterial pathogen leads to debilitating exacerbation of pulmonary infection and constitutes the major cause of morbidity and mortality in CF populations. Colonization of the CF lung by *P. aeruginosa* generally persists despite the use of long-term antibiotic therapy, since antibiotic treatment rarely results in complete
30 eradication of the infection.

As current antibiotic therapies offer limited effectiveness in treating biofilm infection, a need exists for developing therapeutic agents that prevent biofilm formation. The discovery of polypeptides that regulate biofilm formation and polynucleotides encoding such polypeptides fulfills a need in the art by providing new compositions that are useful in the diagnosis, treatment, and prevention of bacterial infection and pathogenesis, as well as biofilm formation in both industrial and medical settings.

Summary of the Invention

As is described in more detail below, we have discovered a regulatory system that modulates microbial phenotypic switching. In one aspect, the invention features an isolated polypeptide that includes an amino acid sequence that is at least 50% (and preferably 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95-99%) identical to the amino acid sequence of PvrR (SEQ ID NO:2), wherein expression of the polypeptide, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism. In preferred embodiments, the polypeptide includes the amino acid sequence of PvrR (SEQ ID NO:2) or consists essentially of the amino acid sequence of PvrR (SEQ ID NO:2) or a fragment thereof.

In a related aspect, the invention features an isolated polypeptide fragment of an isolated polypeptide that includes an amino acid sequence having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2). In preferred embodiments, such a polypeptide fragment includes at least 300 contiguous amino acid residues of the amino acid sequence of PvrR (SEQ ID NO:2). In other embodiments, the fragment is at least 250 amino acid residues, 200 amino acid residues, or 100 amino acid residues of the amino acid sequence of PvrR (SEQ ID NO:2).

In another aspect, the invention features an isolated polynucleotide having at least 50% identity to the nucleotide sequence of *pvrR* (SEQ ID NO:1), wherein expression of the polynucleotide, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism. In preferred embodiments, the isolated polynucleotide includes the nucleotide sequence of *pvrR* (SEQ ID NO:1) or a

complement thereof. In yet other preferred embodiments, the polynucleotide consists essentially of the nucleotide sequence of *pvrR* (SEQ ID NO:1) or a fragment thereof.

In still other related aspects, the invention features a vector including any of the aforementioned isolated polynucleotides and a host cell that includes the vector.

5 The invention further features a variety of screening assays for identifying compounds that modulate phenotype-mediated antibiotic-resistance, biofilm formation, or biofilm-mediated antibiotic resistance. For example, the invention features a screening method that is useful for identifying a compound that modulates the gene expression of a regulator polynucleotide that affects phenotype-mediated antibiotic-
10 resistance in a microorganism. Such a method includes the steps of: (a) providing a microbial cell (e.g., *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*) that includes a polynucleotide having at least 50% identity to the nucleotide sequence of *pvrR* (SEQ ID NO:1)(or a nucleotide sequence that is substantially identical to *pvrR*), wherein expression of the polynucleotide, in the microbial cell, affects phenotype-mediated
15 antibiotic-resistance in the microbial cell; (b) contacting the microbial cell with a compound; and (c) comparing the level of gene expression of the polynucleotide in the presence of the compound with the level of gene expression in the absence of the compound; wherein a measurable difference in gene expression indicates that the compound modulates gene expression of a regulator polynucleotide that affects
20 phenotype-mediated antibiotic-resistance in a microorganism.

In preferred embodiments, the screening method identifies a compound that increases or decreases transcription of the regulator polynucleotide. In other embodiments, the screening method identifies a compound that increases or decreases translation of an mRNA transcribed from the regulator polynucleotide.

25 In other preferred embodiments, the microbial cell is a phenotypic variant (e.g., a small colony variant) having increased biofilm formation. Preferably, the small colony variant is a small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*. In still other embodiments, the small colony variant is a rough small colony variant, for example, a rough small colony variant of *Pseudomonas*, *Vibrio*,

Salmonella, or *Staphylococcus*. In a preferred embodiment, the rough small colony variant is *Pseudomonas aeruginosa* PA14 RSCV.

In other preferred embodiments, the activity of the compound used in the screening assay is dependent upon the presence of the *pvrR* gene (SEQ ID NO:1) or a functional equivalent thereof. For example, the identified compound targets and interacts with the *pvrR* gene (SEQ ID NO:1) or a functional equivalent thereof. In still other preferred embodiments, the expression of the regulator polynucleotide mediates phenotypic switching of the microbial cell in the presence of a high concentration of an antibiotic. In other preferred embodiments of the screening method, the polypeptide is expressed using an isolated polynucleotide that expresses a polypeptide having an amino acid sequence having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2) or a fragment thereof.

In another aspect, the invention features a screening method for identifying a compound that modulates an activity of a polypeptide that affects phenotype-mediated antibiotic-resistance in a microorganism. The method, in general, includes the steps of: (a) providing a microbial cell expressing a polypeptide having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2) (or a polypeptide that is substantially identical to PvrR), wherein expression of the polypeptide, in the microbial cell, affects phenotype-mediated antibiotic-resistance in the microbial cell; (b) contacting the microbial cell with a compound; and (c) comparing an activity of the polypeptide in the presence of the compound with the activity in the absence of the compound; wherein a measurable difference in the activity indicates that the compound modulates the activity of the polypeptide that affects phenotype-mediated antibiotic-resistance in a microorganism. In preferred embodiments, the screening method identifies a compound that increases or decreases the activity of the polypeptide. Comparison of the activity of the polypeptide includes a variety of standard biochemical analyses including immunological assays.

In preferred embodiments, the microbial cell utilized in the screening assay is a phenotypic variant (e.g., *Pseudomonas aeruginosa* PA14 RSCV) having increased biofilm formation relative to wild-type.

In other preferred embodiments, the regulator polypeptide is an isolated polypeptide that includes an amino acid sequence having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2) (or a polypeptide that is substantially identical to PvrR). In particular, such a polypeptide has the ability to regulate

5 phenotypic switching; to regulate biofilm-mediated antibiotic-resistance; to mediate phenotypic switching of the microbial cell in the presence of a high concentration of an antibiotic; or to affect susceptibility of the microbial cell to antibiotic treatment; or to regulate, or mediate, or affect, or any combination of the aforementioned activities thereof. In other preferred embodiments, the regulator polypeptide is an element of a

10 two-component regulatory system. In yet other preferred embodiments, the polypeptide is expressed by an isolated polynucleotide having at least 50% identity to the nucleotide sequence of *pvrR* (SEQ ID NO:1) or a fragment thereof.

Typically, the activity of the compound identified in the screening assay is dependent upon the presence of the PvrR polypeptide (SEQ ID NO:2) or a functional

15 equivalent thereof. In particular aspects of the screening assay, the compound targets the PvrR polypeptide (SEQ ID NO:2) or a functional equivalent thereof.

In another aspect, the invention features a screening method for identifying a compound that modulates microbial biofilm formation. This method, in general, includes the steps of: (a) culturing a microbial cell (e.g., *Pseudomonas*, *Vibrio*,

20 *Salmonella*, or *Staphylococcus*) that includes a polypeptide having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2) (or a polypeptide that is substantially identical to PvrR), wherein the microbial cell, upon culturing, forms a biofilm; (b) contacting the microbial cell with a compound; and (c) comparing microbial biofilm formation in the presence of the compound with microbial biofilm formation in

25 the absence of the compound; wherein a measurable difference in the microbial biofilm formation indicates that the compound modulates biofilm formation.

In preferred embodiments, the screening method identifies a compound that increases or decreases biofilm formation. Typically, such biofilm formation is measured by using any standard method, for example, by assaying microbial aggregation (e.g., by

30 using a microscope); using a salt aggregation test; or by using an attachment assay.

In preferred embodiments, the microbial cell is a phenotypic variant having increased biofilm formation when compared to its wild-type such as a small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*. In other preferred embodiments, the small colony variant is a rough small colony variant of *Pseudomonas*,
5 *Vibrio*, or *Salmonella*. In a preferred embodiment, the rough small colony variant is *Pseudomonas aeruginosa* PA14 RSCV.

In yet other preferred embodiments, the activity of the compound utilized in the screening assay is dependent upon the presence of PvrR polypeptide (SEQ ID NO: 2) or a functional equivalent thereof. For example, the identified compound targets and
10 interacts with the PvrR polypeptide (SEQ ID NO:2) or a functional equivalent thereof, resulting in increasing or decreasing its functional activity.

In still another embodiment, the expression of the polypeptide mediates phenotypic switching of the microbial cell in the presence of a high concentration of an antibiotic.

15 In another embodiment, the polypeptide is an isolated polypeptide that includes an amino acid sequence having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2), wherein expression of the polypeptide, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism.

In still another aspect, the invention features a method of treating a microbial
20 infection involving a microorganism that forms a biofilm in a mammal. The method, in general, includes administering to the mammal a therapeutically-effective amount of a compound that induces or represses expression or activity of a polypeptide that includes an amino acid sequence having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2) (or a polypeptide that is substantially identical to PvrR) or a fragment
25 thereof, wherein expression of the polypeptide or the fragment thereof, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism.

In preferred embodiments, the method further includes administering to the mammal a therapeutically-effective amount of an antibiotic. The treatment is particularly useful for treating patients having cystic fibrosis or a chronic microbial

infection or both. In other preferred embodiments, the microorganism treated using the method belongs to the genus *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*.

In yet another aspect, the invention features a method of cleaning, disinfecting, or decontaminating a surface at least partially covered by a microorganism that forms a
5 biofilm, the method involving contacting the microorganism with a cleaning composition including a compound that induces or represses expression or activity of a polypeptide that includes an amino acid sequence having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2) (or a polypeptide that is substantially identical to PvrR) or fragment thereof, wherein expression of the polypeptide or the
10 fragment thereof, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism.

In yet another aspect, the invention features a screening method for identifying a compound that decreases pathogenicity of an antibiotic-resistant phenotypic variant. The method, in general, includes the steps of: (a) contacting an antibiotic-resistant
15 phenotypic variant with a candidate compound; and (b) measuring reversion of the antibiotic-resistant phenotypic variant to a wild-type phenotype, an increase in reversion indicating that the compound decreases pathogenicity of the antibiotic-resistant phenotypic variant. In preferred embodiments, the antibiotic-resistant phenotypic variant is cultured in the absence of an antibiotic, has increased biofilm formation; is a
20 rough small colony variant; is a hyperpilated variant; has increased hydrophobicity; has an alteration in a surface component; or is a pathogen such as a Gram positive bacterium (e.g., *Staphylococcus*) or a Gram negative bacterium (e.g., *Vibrio*, *Pseudomonas*, or *Salmonella*).

In another aspect, the invention features a screening method for identifying a
25 compound that decreases pathogenicity of an antibiotic-resistant phenotypic variant. The method, in general, includes the steps of: (a) culturing an antibiotic-resistant phenotypic variant with a candidate compound in the presence of an antibiotic; and (b) comparing the number of antibiotic-resistant phenotypic variants in the presence of the compound to the number of antibiotic-resistant phenotypic variants in the absence of the
30 compound, a decrease in the number of the antibiotic-resistant phenotypic variants in the

presence of the compound indicating that the compound decreases pathogenicity of the antibiotic-resistant phenotypic variant.

In yet another aspect, the invention features a screening method for identifying a polynucleotide encoding a regulator polypeptide, the method including the steps of: (a) providing a mutagenized microbe; (b) culturing the mutagenized microbe in the presence of an antibiotic; and (c) comparing the mutagenized microbe with a control wild-type microbe, wherein a change in the number of phenotypic variants identifies the mutagenized microbe as having a mutation in a polynucleotide encoding a regulator polypeptide. In preferred embodiments, the phenotypic variant is a small colony variant.

In another aspect, the invention features a screening method for identifying a polynucleotide encoding a regulator polypeptide that modulates an antibiotic-resistant phenotype of a microorganism. The method, in general, includes the steps of: (a) identifying an antibiotic-resistant phenotypic variant of a microorganism including a first phenotype; (b) mutagenizing the antibiotic-resistant phenotypic variant of the microorganism, thereby generating a mutated phenotypic variant of the microorganism; and (c) selecting the mutated phenotypic variant of step (b) having a second phenotype, other than the first phenotype of the antibiotic-resistant phenotypic variant, wherein the second phenotype identifies a mutation in the mutated phenotypic variant of step (b); and (d) using the mutation for identifying a polynucleotide encoding a regulator polypeptide that modulates an antibiotic-resistant phenotype of a microorganism. In preferred embodiments, the second phenotype includes a wild-type phenotype.

In yet another aspect, the invention features a screening method for identifying a polynucleotide encoding a regulator polypeptide that modulates phenotype-mediated antibiotic-resistance of a microorganism. The method, in general, includes the steps of: (a) transforming an antibiotic-resistant phenotypic variant of a microorganism with a candidate polynucleotide encoding a regulator polypeptide; and (b) culturing the transformed antibiotic-resistant phenotypic variant of a microorganism under conditions suitable for expression of the regulator polypeptide; and (c) measuring reversion of the transformed antibiotic-resistant phenotypic variant of the microorganism to a wild-type

phenotype, an increase in reversion identifies the polynucleotide as encoding a regulator polypeptide that modulates phenotype-mediated antibiotic-resistance.

In preferred embodiments, the polynucleotide encodes a regulator polypeptide that modulates a phenotypic switch from an antibiotic-resistant phenotype to an antibiotic-susceptible phenotype. In other preferred embodiments, the candidate polynucleotide has at least 50% identity to the nucleotide sequence of *pvrR* (SEQ ID NO:1) (or a polynucleotide sequence that is substantially identical to *pvrR*). In other embodiments, the candidate polynucleotide sequence is substantially identical to any one of the polynucleotides shown in Figures 5B, 5C, 6A-6K, and 7A-7E. In other preferred
5
10
embodiments, the candidate polynucleotide encodes a polypeptide that is an element of a two-component regulatory system.

In another aspect, the invention features an isolated polypeptide including an amino acid sequence that is substantially identical to the amino acid sequence of any one the polypeptides shown in Figures 5E (SEQ ID NO: 4) and 6L-6V (SEQ ID NOS: 19-
15
29), each of which are encoded by a polynucleotide of the ORF1 region.

For example, with respect to the ORF1 region, the invention features an isolated polypeptide that includes an amino acid sequence that is at least 50% (and preferably 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95-99%) identical to the amino acid sequence of the polypeptide shown in Figure 5E (SEQ ID NO: 4) or to a polypeptide shown in Figures 6L-6V (SEQ ID NOS: 19-29), wherein expression of the polypeptide, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism. Preferably, the polypeptide includes the amino acid sequence shown in Figure 5E or consists essentially of the amino acid sequence shown in Figure 5E or a fragment thereof.
20

In a related aspect, the invention features an isolated polypeptide fragment of an isolated polypeptide that includes an amino acid sequence having at least 50% identity to the amino acid sequence the polypeptide shown in Figure 5E or to a polypeptide shown in any one of Figures 6L-6V. In preferred embodiments, such a polypeptide fragment includes at least 400 contiguous amino acid residues of the amino acid sequence shown
25
30
in any one of Figures 5E and 6L-6V. In other embodiments, the fragment is at least 300

amino acid residues, 200 amino acid residues, or 100 amino acid residues of the polypeptides shown in Figures 5E and 6L-6V.

In another aspect, the invention features an isolated polynucleotide molecule including a sequence substantially identical to any one of the polynucleotides shown in
5 Figures 5B (SEQ ID NO:3) and 6A-6K (SEQ ID NOS: 8-18), which are found in the ORF1 region. In preferred embodiments, the isolated polynucleotide molecule has at least 45%, 50%, 60%, 70%, 80%, 90%, or even 95-99% identity to any one of these isolated molecules.

For example, with respect to the ORF1 region, the invention features an isolated
10 polynucleotide having at least 50% identity to the nucleotide sequence shown in Figure 5B or to any one of the nucleotide sequences shown in Figures 6A-6K, wherein expression of the polynucleotide, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism. In preferred embodiments, the isolated polynucleotide includes the nucleotide sequence shown in Figure 5B or a complement
15 thereof. In yet other preferred embodiments, the polynucleotide consists essentially of the nucleotide sequence shown in Figure 5B or a fragment thereof.

In still other related aspects, the invention features a vector including any of the aforementioned isolated polynucleotides and a host cell that includes the vector.

The invention further features a variety of screening assays for identifying
20 compounds that modulate phenotype-mediated antibiotic-resistance, biofilm formation, or biofilm-mediated antibiotic resistance. For example, the invention features a screening method that is useful for identifying a compound that modulates the gene expression of a regulator polynucleotide that affects phenotype-mediated antibiotic-resistance in a microorganism. Such a method includes the steps of: (a) providing a
25 microbial cell (e.g., *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*) that includes a polynucleotide that is substantially identical to any one of the nucleotide sequences shown in Figures 5B or 6A-6K (or a polynucleotide having at least 40% identity to any one of these sequences), wherein expression of the polynucleotide, in the microbial cell, affects phenotype-mediated antibiotic-resistance in the microbial cell; (b) contacting the
30 microbial cell with a compound; and (c) comparing the level of gene expression of the

polynucleotide in the presence of the compound with the level of gene expression in the absence of the compound; wherein a measurable difference in gene expression indicates that the compound modulates gene expression of a regulator polynucleotide that affects phenotype-mediated antibiotic-resistance in a microorganism.

5 In preferred embodiments, the screening method identifies a compound that increases or decreases transcription of the regulator polynucleotide. In other embodiments, the screening method identifies a compound that increases or decreases translation of an mRNA transcribed from the regulator polynucleotide.

 In other preferred embodiments, the microbial cell is a phenotypic variant (e.g., a
10 small colony variant) having increased biofilm formation. Preferably, the small colony variant is a small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*. In still other embodiments, the small colony variant is a rough small colony variant, for example, a rough small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*. In a preferred embodiment, the rough small colony
15 variant is *Pseudomonas aeruginosa* PA14 RSCV.

 In other preferred embodiments, the activity of the compound used in the screening assay is dependent upon the presence of any one of the polynucleotides shown in Figures 5B or 6A-6K, or a functional equivalent thereof. For example, the identified compound targets any one of the polynucleotides shown in Figures 5B or 6A-6K or a
20 functional equivalent thereof. In still other preferred embodiments, the expression of the regulator polynucleotide mediates phenotypic switching of the microbial cell in the presence of a high concentration of an antibiotic. In other preferred embodiments of the screening method, the polypeptide is expressed using an isolated polynucleotide that encodes a polypeptide that is substantially identical to any one of the polynucleotides
25 shown Figures 5B and 6A-6K or a fragment thereof.

 In another aspect, the invention features a screening method for identifying a compound that modulates an activity of a polypeptide that affects phenotype-mediated antibiotic-resistance in a microorganism. The method, in general, includes the steps of:
(a) providing a microbial cell expressing a polypeptide that is substantially identical to
30 any one of the polypeptides shown in Figures 5E and 6L-6V (or a polypeptide having at

least 40% identity to any one of these sequences), wherein expression of the polypeptide, in the microbial cell, affects phenotype-mediated antibiotic-resistance in the microbial cell; (b) contacting the microbial cell with a compound; and (c) comparing an activity of the polypeptide in the presence of the compound with the activity in the
5 absence of the compound; wherein a measurable difference in the activity indicates that the compound modulates the activity of the polypeptide that affects phenotype-mediated antibiotic-resistance in a microorganism. In preferred embodiments, the screening method identifies a compound that increases or decreases the activity of the polypeptide. Comparison of the activity of the polypeptide includes a variety of standard biochemical
10 analyses including immunological assays.

In preferred embodiments, the microbial cell utilized in the screening assay is a phenotypic variant (e.g., *Pseudomonas aeruginosa* PA14 RSCV) having increased biofilm formation.

In other preferred embodiments, the regulator polypeptide is an isolated
15 polypeptide that includes an amino acid sequence that is substantially identical to any one of the polypeptides shown in Figures 5E and 6L-6V (or a polypeptide having at least 40% identity to any one of these sequences). In particular, such a polypeptide has the ability to regulate phenotypic switching; to regulate biofilm-mediated antibiotic-resistance; to mediate phenotypic switching of the microbial cell in the presence of a
20 high concentration of an antibiotic; or to affect susceptibility of the microbial cell to antibiotic treatment; or any combination thereof. In other preferred embodiments, the regulator polypeptide is an element of a two-component regulatory system. In yet other preferred embodiments, the polypeptide is expressed by an isolated polynucleotide that is substantially identical to any one of the nucleotide sequences shown in Figures 5B and
25 6A-6K (or a polynucleotide having at least 40% identity to any one of these sequences) or a fragment thereof, upon which the activity of the regulator polypeptide is increased or decreased.

Typically, the activity of the compound identified in the screening assay is dependent upon the presence of any one of the polypeptides shown in Figures 5E and
30 6L-6V or a functional equivalent thereof. In particular aspects of the screening assay,

the compound targets or interacts with any one of the polypeptides shown in Figures 5E and 6L-6V or a functional equivalent thereof.

In another aspect, the invention features a screening method for identifying a compound that modulates microbial biofilm formation. This method, in general, includes the steps of: (a) culturing a microbial cell (e.g., *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*) that includes a polypeptide that is substantially identical to any one of the polypeptides shown in Figures 5E and 6L-6V (or a polypeptide having at least 40% identity to any one of these sequences), wherein the microbial cell, upon culturing, forms a biofilm; (b) contacting the microbial cell with a compound; and (c) comparing microbial biofilm formation in the presence of the compound with microbial biofilm formation in the absence of the compound; wherein a measurable difference in the microbial biofilm formation indicates that the compound modulates biofilm formation.

In preferred embodiments, the screening method identifies a compound that increases or decreases biofilm formation. Typically, such biofilm formation is measured by using any standard method, for example, by assaying microbial aggregation (e.g., by using a microscope); using a salt aggregation test; or by using an attachment assay.

In preferred embodiments, the microbial cell is a phenotypic variant having increased biofilm formation when compared to its wild-type such as a small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*. In other preferred embodiments, the small colony variant is a rough small colony variant of *Pseudomonas*, *Vibrio*, or *Salmonella*.

In yet other preferred embodiments, the activity of the compound utilized in the screening assay is dependent upon the presence of the polypeptide or a functional equivalent thereof. For example, the identified compound targets or interacts with the polypeptide or a functional equivalent thereof, resulting in increasing or decreasing its functional activity.

In still another embodiment, the expression of the polypeptide mediates phenotypic switching of the microbial cell in the presence of a high concentration of an antibiotic.

In another embodiment, the polypeptide is an isolated polypeptide that includes an amino acid sequence that is substantially identical to any one of the polypeptides shown in Figures 5E and 6L-6V (or a polypeptide having at least 40% identity to any one of these sequences), wherein expression of the polypeptide, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism.

In still another aspect, the invention features a method of treating a microbial infection involving a microorganism that forms a biofilm in a mammal. The method, in general, includes administering to the mammal a therapeutically-effective amount of a compound that induces or represses expression or activity of a polypeptide that includes a polypeptide that is substantially identical to any one of the polypeptides shown in Figures 5E and 6L-6V or a fragment thereof (or a polypeptide having at least 40% identity to any one of these sequences), wherein expression of the polypeptide or the fragment thereof, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism.

In another aspect, the invention features an isolated polypeptide including an amino acid sequence that is substantially identical to the amino acid sequence of any one of the polypeptides shown in Figures 5F and Figures 7F-7J, each of which are encoded by a polynucleotide of the ORF3 region.

For example, with respect to the ORF3 region, the invention features an isolated polypeptide that includes an amino acid sequence that is at least 50% (and preferably 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95-99%) identical to the amino acid sequence of any one of the polypeptides shown in Figures 5F (SEQ ID NO:6) and 7F-7J (SEQ ID NOS:35-39), wherein expression of the polypeptide, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism. Preferably, the polypeptide includes the amino acid sequence shown in Figure 7J (SEQ ID NO:39) or consists essentially of the amino acid sequence shown in Figures 5F (SEQ ID NO:6) and 7F-7I (SEQ ID NOS:35-38) or a fragment thereof.

In a related aspect, the invention features an isolated polypeptide fragment of an isolated polypeptide that includes an amino acid sequence having at least 50% identity to the amino acid sequence of the polypeptides shown in Figures 5F and 7F-7J. In

preferred embodiments, such a polypeptide fragment includes at least 300 contiguous amino acid residues of the amino acid sequence shown in any one of Figures 5F and 7F-7J. In other embodiments, the fragment is at least 200 amino acid residues, or 100 amino acid residues of the polypeptides shown in Figures 5F and 7F-7J.

5 In another aspect the invention features an isolated polynucleotide molecule including a sequence substantially identical to any one of the polynucleotides shown in Figures 5C (SEQ ID NO:5) and 7A-7E (SEQ ID NOS:30-34). In preferred embodiments, the isolated polynucleotide molecule has at least 45%, 50%, 60%, 70%, 80%, 90%, or even 95% identity to any one of these molecules.

10 For example with respect to the ORF3 region, the invention features an isolated polynucleotide having at least 50% identity to any one of the nucleotide sequences shown in Figures 5C and 7A-7E, wherein expression of the polynucleotide, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism. In preferred embodiments, the isolated polynucleotide includes the nucleotide sequence
15 shown in Figure 5C or a complement thereof. In yet other preferred embodiments, the polynucleotide consists essentially of the nucleotide sequence shown in Figure 5C or a fragment thereof.

 In still other related aspects, the invention features a vector including any of the aforementioned isolated polynucleotides and a host cell that includes the vector.

20 The invention further features a variety of screening assays for identifying compounds that modulate phenotype-mediated antibiotic-resistance, biofilm formation, or biofilm-mediated antibiotic resistance. For example, the invention features a screening method that is useful for identifying a compound that modulates the gene expression of a regulator polynucleotide that affects phenotype-mediated antibiotic-
25 resistance in a microorganism. Such a method includes the steps of: (a) providing a microbial cell (e.g., *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*) that includes a polynucleotide substantially identical to the nucleotide sequences shown in Figures 5C and 7A-7E (or a polynucleotide having at least 45% identity to any one of these sequences), wherein expression of the polynucleotide, in the microbial cell, affects
30 phenotype-mediated antibiotic-resistance in the microbial cell; (b) contacting the

microbial cell with a compound; and (c) comparing the level of gene expression of the polynucleotide in the presence of the compound with the level of gene expression in the absence of the compound; wherein a measurable difference in gene expression indicates that the compound modulates gene expression of a regulator polynucleotide that affects phenotype-mediated antibiotic-resistance in a microorganism.

In preferred embodiments, the screening method identifies a compound that increases or decreases transcription of the regulator polynucleotide. In other embodiments, the screening method identifies a compound that increases or decreases translation of an mRNA transcribed from the regulator polynucleotide.

In other preferred embodiments, the microbial cell is a phenotypic variant (e.g., a small colony variant) having increased biofilm formation. Preferably, the small colony variant is a small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*. In still other embodiments, the small colony variant is a rough small colony variant, for example, a rough small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*. In a preferred embodiment, the rough small colony variant is *Pseudomonas aeruginosa* PA14 RSCV.

In other preferred embodiments, the activity of the compound used in the screening assay is dependent upon the presence of any one of the polynucleotides shown in Figures 5C and 7A-7E or a functional equivalent thereof. For example, the identified compound targets or interacts with any one of the polynucleotides shown in Figures 5C and 7A-7E or a functional equivalent thereof. In still other preferred embodiments, the expression of the regulator polynucleotide mediates phenotypic switching of the microbial cell in the presence of a high concentration of an antibiotic. In other preferred embodiments of the screening method, the polypeptide is expressed from an isolated polynucleotide that expresses a polypeptide that includes an amino acid sequence having at least 50% identity to any one of the amino acid sequences shown in Figures 5F and 7F-7J or a fragment thereof.

In another aspect, the invention features a screening method for identifying a compound that modulates an activity of a polypeptide that affects phenotype-mediated antibiotic-resistance in a microorganism. The method, in general, includes the steps of:

(a) providing a microbial cell expressing a polypeptide that is substantially identical to any one of the polypeptides shown in Figures 5F and 7F-7J (or a polypeptide having at least 45% identity to any one of these sequences), wherein expression of the polypeptide, in the microbial cell, affects phenotype-mediated antibiotic-resistance in the microbial cell; (b) contacting the microbial cell with a compound; and (c) comparing an activity of the polypeptide in the presence of the compound with the activity in the absence of the compound; wherein a measurable difference in the activity indicates that the compound modulates the activity of the polypeptide that affects phenotype-mediated antibiotic-resistance in a microorganism. In preferred embodiments, the screening method identifies a compound that increases or decreases the activity of the polypeptide. Comparison of the activity of the polypeptide includes a variety of standard biochemical analyses including immunological assays.

In preferred embodiments, the microbial cell utilized in the screening assay is a phenotypic variant (e.g., *Pseudomonas aeruginosa* PA14 RSCV) having increased biofilm formation.

In other preferred embodiments, the regulator polypeptide is an isolated polypeptide that includes an amino acid sequence that is substantially identical to any one of the polypeptides shown in Figures 5F and 7F-7J (or a polypeptide having at least 45% identity to any one of these sequences). In particular, such a polypeptide has the ability to regulate phenotypic switching; to regulate biofilm-mediated antibiotic-resistance; to mediate phenotypic switching of the microbial cell in the presence of a high concentration of an antibiotic; or to affect susceptibility of the microbial cell to antibiotic treatment; or any combination thereof. In other preferred embodiments, the regulator polypeptide is an element of a two-component regulatory system. In yet other preferred embodiments, the polypeptide is expressed by an isolated polynucleotide substantially identical to any one of the nucleotide sequences shown in Figures 5C and 7A-7E (or by a polynucleotide having at least 45% identity to any one of these sequences) or a fragment thereof, upon which the activity of the regulator polypeptide is increased or decreased.

Typically, the activity of the compound identified in the screening assay is dependent upon the presence of any one of the polypeptides shown in Figures 5F and 7F-7J or a functional equivalent thereof. In particular aspects of the screening assay, the compound targets and interacts with the polypeptide of Figures 5F and 7F-7J or a functional equivalent thereof.

In another aspect, the invention features a screening method for identifying a compound that modulates microbial biofilm formation. This method, in general, includes the steps of: (a) culturing a microbial cell (e.g., *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*) that includes a polypeptide substantially identical to any one of the amino acid sequences shown in Figures 5F and 7F-7J (or a polypeptide having at least 45% identity to any one of these sequences), wherein the microbial cell, upon culturing, forms a biofilm; (b) contacting the microbial cell with a compound; and (c) comparing microbial biofilm formation in the presence of the compound with microbial biofilm formation in the absence of the compound; wherein a measurable difference in the microbial biofilm formation indicates that the compound modulates biofilm formation.

In preferred embodiments, the screening method identifies a compound that increases or decreases biofilm formation. Typically, such biofilm formation is measured by using any standard method, for example, by assaying microbial aggregation (e.g., by using a microscope); using a salt aggregation test; or by using an attachment assay.

In preferred embodiments, the microbial cell is a phenotypic variant having increased biofilm formation when compared to its wild-type such as a small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*. In other preferred embodiments, the small colony variant is a rough small colony variant of *Pseudomonas*, *Vibrio*, or *Salmonella*.

In yet other preferred embodiments, the activity of the compound utilized in the screening assay is dependent upon the presence of the polypeptide or a functional equivalent thereof. For example, the identified compound targets and interacts with the polypeptide or a functional equivalent thereof, resulting in increasing or decreasing its functional activity.

In still another embodiment, the expression of the polypeptide mediates phenotypic switching of the microbial cell in the presence of a high concentration of an antibiotic.

5 In another embodiment, the polypeptide is an isolated polypeptide that includes an amino acid sequence that is substantially identical to any one of the amino acid sequences shown in Figures 5F and 7F-7J (or a polypeptide having at least 45% identity to any one of these sequences), wherein expression of the polypeptide, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism.

10 In still another aspect, the invention features a method of treating a microbial infection involving a microorganism that forms a biofilm in a mammal. The method, in general, includes administering to the mammal a therapeutically-effective amount of a compound that induces or represses expression or activity of a polypeptide that includes an amino acid sequence that is substantially identical to any one of the amino acid sequences shown in Figures 5F and 7F-7J or a fragment thereof (or a polypeptide having
15 at least 45% identity to any one of these sequences), wherein expression of the polypeptide or the fragment thereof, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism.

In preferred embodiments, the method further includes administering to the mammal a therapeutically-effective amount of an antibiotic. The treatment is
20 particularly useful for treating patients having cystic fibrosis or a chronic infection or both. In other preferred embodiments, the microorganism treated using the method belongs to the genus *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*.

In yet another aspect, the invention features a method of cleaning, disinfecting, or decontaminating a surface at least partially covered by a microorganism that forms a
25 biofilm, the method involving contacting the microorganism with a cleaning composition including a compound that induces or represses expression or activity of a polypeptide that includes an amino acid sequence having at least 50% identity to the amino acid sequence of Figures 5E, 5F, 6L-6V, and 7F-7J or fragment thereof (or a polypeptide that is substantially identical to any one of these polypeptides), wherein

expression of the polypeptide or the fragment thereof, in a microorganism, affects phenotype-mediated antibiotic-resistance in the microorganism.

The invention also features methods for identifying compounds useful for treating a patient having a biofilm infection. The method includes the steps of

5 contacting a biofilm *in vitro* with (i) an antibiotic and (ii) a candidate compound (e.g., a compound that modulates the expression, at the transcriptional, post-transcriptional, translational, or post-translational levels, of a polynucleotide having at least 50% identity to any of the polynucleotides described herein (or that is substantially identical to a polynucleotide described herein), and determining whether the biofilm grows more

10 slowly than (a) biofilm cells contacted with an antibiotic but not contacted with the test compound, and (b) biofilm cells contacted with the candidate compound but not with the antibiotic. In another embodiment, the biofilm is contacted with two or more different antibiotics. Exemplary antibiotics useful in the method include amikacin, aminoglycosides (e.g., tobramycin), aztreonam, carbenicillin, cephalosporines (e.g.,

15 ceftazidime or cefipime), chloramphenicol, gentamicin, levofloxacin, meropenem, piperacillin, tazobactam, tetracycline, and quinolones (e.g., ciprofloxacin). A candidate compound that reduces biofilm formation in the presence of an antibiotic (or combination of different antibiotics), but does not decrease biofilm formation in the absence of the antibiotic (or combination of different antibiotics), is a compound that is

20 useful in combination therapy for treating a patient having a biofilm infection.

The invention further features a method for treating a patient having a biofilm infection, by administering to the patient an antibiofilm combination therapy that includes a compound identified as modulating expression, at the transcriptional, post-transcriptional, translational, or post-translational levels, of a polynucleotide having at

25 least 50% identity to any of the polynucleotides described herein (or that is substantially identical to a polynucleotide described herein) and one or more antibiotics, including, but not limited to, amikacin, aminoglycosides (e.g., tobramycin), aztreonam, carbenicillin, cephalosporines (e.g., ceftazidime or cefipime), chloramphenicol, gentamicin, levofloxacin, meropenem, piperacillin, tazobactam, tetracycline, and

quinolones (e.g., ciprofloxacin), simultaneously or within a period of time (e.g., 14 to 21 days) sufficient to inhibit the growth of the biofilm.

Preferably, the compound and antibiotic are administered within fifteen days of each other, more preferably within five or ten days of each other, and most preferably within twenty-four hours of each other or even simultaneously. Exemplary biofilms treated according to any of the methods described herein are those formed by bacteria, including but not limited to, *Pseudomonas*, *Staphylococcus*, *Salmonella*, *Vibrio*, *Haemophilus*, *Mycobacterium*, *Helicobacter*, *Burkholderia*, or *Streptococci*.

In a related aspect, the invention also features a method for treating a patient having a biofilm such as one formed from *Pseudomonas* (e.g., *Pseudomonas aeruginosa*). In this method, a patient is administered (a) a first compound (e.g., a compound that modulates the expression, at the transcriptional, post-transcriptional, translational, or post-translational; of a polynucleotide having at least 50% identity to a polynucleotide described herein (or that is substantially identical to a polynucleotide described herein)), and (b) one or more antibiotics (such as amikacin, aminoglycosides (e.g., tobramycin), aztreonam, carbenicillin, cephalosporines (e.g., ceftazidime or cefipime), chloramphenicol, gentamicin, levofloxacin, meropenem, piperacillin, tazobactam, tetracycline, and quinolones (e.g., ciprofloxacin). If desired, the therapy includes administration of two antibiotics according to standard methods known in the art. Such dual antibiotic combinations most preferably include high-dose tobramycin plus meropenem, meropenem plus ciprofloxacin, or tobramycin (4 µg/ml), or cefipime. Other preferred combinations include piperacillin plus tazobactam, or piperacillin plus ciprofloxacin. The antibiotic and compound combination therapy are preferably administered simultaneously or within a period of time sufficient to inhibit the growth of the biofilm.

In any of the foregoing treatments, the compound and antibiotic included in the combination therapy are preferably administered to the patient as part of a pharmaceutical composition that also includes a pharmaceutically acceptable carrier. Preferred modes of administration include intramuscular, intravenous, inhalation, and oral administration, or a combination thereof.

The antibiofilm combinations of the invention can also be part of a pharmaceutical kit. Preferably, the first compound (e.g., a compound identified as modulating expression, at the transcriptional, post-transcriptional, translational, or post-translational levels, of a polynucleotide or polypeptide having at least 50% identity to
5 any one of the polynucleotide or polypeptide sequences described herein (or that is substantially identical to any one of the polynucleotides or polypeptides described herein)) and the second compound, an antibiotic, are formulated together or separately and in individual dosage amounts.

Combination therapy may be provided wherever antibiotic treatment is
10 performed: at home, the doctor's office, a clinic, a hospital's outpatient department, or a hospital. Treatment generally begins at a hospital so that the doctor can observe the therapy's effects closely and make any adjustments that are needed. The duration of the combination therapy depends on the kind of biofilm being treated, the age and condition of the patient, the stage and type of the patient's biofilm infection, and how the patient's
15 body responds to the treatment. Drug administration may be performed at different intervals (e.g., daily, weekly, or monthly) and the administration of each agent can be determined individually. Combination therapy may be given in on-and-off cycles that include rest periods so that the patient's body has a chance to build healthy new cells and regain its strength.

20 By "isolated polynucleotide" is meant a nucleic acid (e.g., a DNA) that is free of the genes which, in the naturally-occurring genome of the organism from which the nucleic acid molecule of the invention is derived, flank the gene. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or
25 eukaryote; or that exists as a separate molecule (for example, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. In addition, the term includes an RNA molecule which is transcribed from a DNA molecule, as well as a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

By "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (for example, glycosylation or phosphorylation).

By an "isolated polypeptide" is meant a polypeptide of the invention that has been separated from components which naturally accompany it. Typically, the polypeptide is isolated when it is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, a polypeptide of the invention. An isolated polypeptide of the invention may be obtained, for example, by extraction from a natural source (for example, a pathogen); by expression of a recombinant nucleic acid encoding such a polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, for example, column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

By "substantially identical" is meant a polypeptide or nucleic acid molecule (e.g., a polynucleotide) exhibiting at least 50% identity to a reference amino acid sequence (for example, any one of the amino acid sequences described herein) or nucleic acid sequence (for example, any one of the nucleic acid sequences described herein). Preferably, such a sequence is at least 60%, more preferably 80%, and most preferably 90% or even 95% identical at the amino acid level or nucleic acid to the sequence used for comparison.

Sequence identity is typically measured using sequence analysis software (for example, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705, BLAST, BESTFIT, GAP, or PILEUP/PRETTYBOX programs). Such software matches identical or similar sequences by assigning degrees of homology to various substitutions, deletions, and/or other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. In an exemplary approach to determining the

degree of identity, a BLAST program may be used, with a probability score between e^{-3} and e^{-100} indicating a closely related sequence.

By “transformed cell” is meant a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a polynucleotide molecule encoding (as used herein) a polypeptide of the invention.

By “positioned for expression” is meant that the polynucleotide of the invention (e.g., a DNA molecule) is positioned adjacent to a DNA sequence which directs transcription and translation of the sequence (i.e., facilitates the production of, for example, a recombinant polypeptide of the invention, or an RNA molecule).

By “purified antibody” is meant an antibody which is at least 60%, by weight, free from proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably 90%, and most preferably at least 99%, by weight, antibody. A purified antibody of the invention may be obtained, for example, by affinity chromatography using a recombinantly-produced polypeptide of the invention and standard techniques.

By “specifically binds” is meant a compound or antibody which recognizes and binds a polypeptide of the invention but which does not substantially recognize and bind other molecules in a sample, for example, a biological sample, which naturally includes a polypeptide of the invention.

By “derived from” is meant isolated from or having the sequence of a naturally-occurring sequence (e.g., a cDNA, genomic DNA, synthetic, or combination thereof).

By “inhibiting biofilm formation” is meant the ability of a candidate compound to decrease the development or progression of biofilm formation. Preferably, such inhibition decreases biofilm formation by at least 1% to 5%, more preferably by at least 10%, 15%, 20%, or 25%, and most preferably by at least 30% to 50%, as compared to biofilm formation in the absence of the candidate compound in any appropriate pathogenicity assay (for example, those assays described herein). In one particular example, inhibition is measured by continuous culture conditions of a microbe exposed to a candidate compound or extract, a decrease in the level of biofilm formation relative

to the level of biofilm formation of the microbe not exposed to the compound indicating compound-mediated inhibition of biofilm formation.

By "biofilm regulator polynucleotide" is meant a polynucleotide encoding a cellular component (e.g., PvrR) that modulates phenotypic switching, such as a phenotypic switch that occurs during biofilm formation, disintegration, or both.

By "phenotypic switching" is meant the reversible alteration of one or more phenotypic characteristics. Such an alteration typically occurs, for example, when a wild-type microbe develops into an antibiotic-resistant phenotypic variant or when an antibiotic-resistant phenotypic variant develops into a wild-type microbe.

By "immunological assay" is meant an assay that relies on an immunological reaction, for example, antibody binding to an antigen. Examples of immunological assays include ELISAs, Western blots, immunoprecipitations, and other assays known to the skilled artisan.

By a "two-component regulatory system" is meant a regulatory system that includes at least two components such as a sensor that senses an environmental signal and a response regulator that modulates one or more effectors.

By "aggregation" is meant a collection of two or more individual microorganisms into a mass or clump, such that the individuals form an aggregated microbial unit. Aggregation can be measured using assays provided herein. Exemplary assays include visual inspection, measuring attachment to a surface, or by assaying for biofilm formation using methods known to the skilled artisan.

By "pathogenicity" is meant the ability of a microorganism to cause disease. A microorganism that forms a biofilm, has increased antibiotic resistance, or displays phenotypic variation is more pathogenic than a wild-type microorganism in that it is less susceptible to conventional antibiotic treatment.

The invention provides a number of targets that are useful for the development of drugs that specifically block the biofilm formation of a microbe. In addition, the methods of the invention provide a facile means to identify compounds that are safe for use in eukaryotic host organisms (i.e., compounds which do not adversely affect the normal development and physiology of the organism), and efficacious against

pathogenic microbes (i.e., by suppressing the virulence of a pathogen). In addition, the methods of the invention provide a route for analyzing virtually any number of compounds for an anti-virulence effect with high-volume throughput, high sensitivity, and low complexity. The methods are also relatively inexpensive to perform and enable the analysis of small quantities of active substances found in either purified or crude
5 extract form.

Other features and advantages of the invention will be apparent from the detailed description, and from the claims.

10 Brief Description of the Drawings

Figure 1A shows the reversion of PA14 rough small colony variants (RSCV) to the wild-type phenotype as observed at the edges of the colonies (arrow) after 2-3 days incubation on antibiotic free LB agar at room temperature.

Figure 1B shows a confocal scanning laser microscopic analysis of bacterial
15 aggregates (arrows) formed by wild-type PA14 and PA14 RSCV expressing green fluorescent protein (GFP) after overnight growth in liquid broth. Scale bar, 25 μm .

Figure 1C shows the attachment of wild-type PA14 and antibiotic resistant variants to polyvinylchloride plastic (PVC) after 6 hours of growth.

Figure 1D shows a confocal laser scanning microscope analysis of biofilm
20 formed by wild-type PA14 and PA14 RSCV expressing GFP in flow-chambers under continuous culture conditions. Scale bar, 50 μm .

Figure 1E shows PA14 and PA14 RSCV biofilm resistance to tobramycin as determined by measuring viable biomass on 45 hour-old established biofilms before (filled bars) and after (open bars) 36-hour tobramycin (200 $\mu\text{g/ml}$) treatment.

25 Figure 2A shows the effect of different environmental stimuli on the rate of appearance of antibiotic resistant variants. This was determined by growing the cultures of wild-type PA14 under the specified conditions on media containing 200 $\mu\text{g/ml}$ kanamycin.

Figure 2B shows the minimal inhibitory concentrations of kanamycin for strain
30 PA14 using the different conditions specified.

Figure 3A shows the reversion of PA14 RSCV present in sputum samples of a cystic fibrosis patient (designated "CF 5") as observed on the edges of the variant colonies (arrow) after prolonged incubation on antibiotic-free medium at room temperature.

5 Figure 3B shows the increased attachment to PVC plastic of antibiotic resistant variants SCV 42 and SCV 43 obtained after plating CF isolates CF 42 and CF 43 on tobramycin (10 µg/ml).

Figure 4A shows the attachment to PVC plastic of PA14, antibiotic resistant variants, and PA14 RSCV carrying pEd202 (PA14 RSCV /pED202) or pUCP19 (PA14 RSCV /pUCP19) after 4 hours of growth was quantitated.

Figure 4B shows the predicted amino acid sequence alignment of PvrR with the sequences that correspond to VieA from *V. Cholerae* and the *P. aeruginosa* PAO1 putative response regulator PA3947 (PAO1 RR). Numbers above the scale indicate number of amino acids. Lower panel contains domain family numbers according to ProDom nomenclature.

Figure 4C shows that the *pvrR* gene is flanked by two open reading frame regions (ORFs), designated *ORF1* and *ORF3*, with the same transcriptional orientation. Start codons within ORFs were assigned based on visual inspection for appropriately spaced ribosome-binding sequences.

20 Figure 4D shows the number of variants resistant to kanamycin (200 µg/ml). This was evaluated after plating overnight cultures of PA14 and PA14 overexpressing PvrR (PA14/pED202).

Figure 4E shows the attachment to PVC plastic of PA14 and PA14 overexpressing PvrR (PA14/pED202) after 12 hours of growth, quantitated as described herein.

Figure 4F shows the number of antibiotic resistant variants for PA14 and the *pvrR* mutant ($\Delta pvrR$) as determined by plating overnight cultures on LB agar containing kanamycin (200 µg/ml).

Figure 5A shows the nucleic acid sequence of *pvrR* (SEQ ID NO:1).

Figure 5B shows the nucleic acid sequence of an *ORF1* polynucleotide (SEQ ID NO:3). This polynucleotide sequence begins at nucleotide 1504 and ends at nucleotide 2919 of SEQ ID NO: 7 as shown in Figure 5G.

Figure 5C shows the nucleic acid sequence of an ORF3 polynucleotide (SEQ ID NO:5). This polynucleotide sequence begins at nucleotide 4385 and ends at nucleotide 6379 of SEQ ID NO:7 as shown in Figure 5G.

Figure 5D shows the deduced amino acid sequence of PvrR (SEQ ID NO:2).

Figure 5E shows the deduced amino acid sequence of a polypeptide (SEQ ID NO:4) encoded by the polynucleotide shown in Figure 5B.

Figure 5F shows the deduced amino acid sequence of a polypeptide (SEQ ID NO:6) encoded by the polynucleotide shown in Figure 5C.

Figure 5G shows the nucleic acid sequence (SEQ ID NO:7) that includes the pvrR gene (SEQ ID NO:1), and the ORF1 (SEQ ID NOS:3 and 8-18) and ORF3 (SEQ ID NOS:5 and 30-34) regions. The start and stop codons for the identified open reading frames are highlighted.

Figures 6A-6K show the nucleotide sequences of several open reading frames identified in the ORF1 region (SEQ ID NO:8 begins at nucleotide 124 and ends at nucleotide 2919; SEQ ID NO:9 begins at nucleotide 199 and ends at nucleotide 2919; SEQ ID NO:10 begins at nucleotide 217 and ends at nucleotide 2919; SEQ ID NO:11 begins at nucleotide 256 and ends at nucleotide 2919; SEQ ID NO:12 begins at nucleotide 295 and ends at nucleotide 2919; SEQ ID NO:13 begins at nucleotide 307 and ends at nucleotide 2919; SEQ ID NO:14 begins at nucleotide 511 and ends at nucleotide 2919; SEQ ID NO:15 begins at nucleotide 760 and ends at nucleotide 2919; SEQ ID NO:16 begins at nucleotide 790 and ends at nucleotide 2919; SEQ ID NO:17 begins at nucleotide 919 and ends at nucleotide 2919; and SEQ ID NO:18 begins at nucleotide 1429 and ends at nucleotide 2919).

Figures 6L-6V show the deduced amino acid sequences of the polypeptides (SEQ ID NOS: 19-29) identified in Figures 6A-6K above.

Figures 7A-7E show the nucleotide sequence of several open reading frames identified in the ORF3 region (SEQ ID NO:30 begins at nucleotide 4388 and ends at

nucleotide 6379; SEQ ID NO:31 begins at nucleotide 4550 and ends at nucleotide 6379; SEQ ID NO:32 begins at nucleotide 4572 and ends at nucleotide 6379; SEQ ID NO:33 begins at nucleotide 4880 and ends at nucleotide 6379; and SEQ ID NO:34 begins at nucleotide 5258 and ends at nucleotide 6379).

- 5 Figures 7F-7J show the deduced amino acid sequences of the polypeptides (SEQ ID NOS:35-39) identified in Figures 7A-7E above.

Detailed Description

Overview

- 10 *Pseudomonas aeruginosa* is the most important pathogen in the lungs of cystic fibrosis (CF) patients. Colonization of the CF lung by *P. aeruginosa* persists despite the use of long-term antibiotic therapy, since antibiotic treatment rarely results in eradication of the infection. Reports have suggested a direct link between resistance to antimicrobial compounds and the ability of *P. aeruginosa* to form biofilm in CF lungs.
- 15 Other hypotheses explain *P. aeruginosa* antibiotic resistance by postulating that factors within the CF respiratory tract select for phenotypic variants suited to survive antimicrobial treatment. As is discussed below, we have determined that a clinical isolate of *P. aeruginosa*, strain PA14, was capable of growing under inhibitory concentrations of the antibiotic kanamycin (up to 40 times the susceptibility level of the
- 20 strain) when bacteria had undergone phenotypic variation. The antibiotic resistant variant colonies obtained from kanamycin plates were smaller in size and had a different colony morphology compared to the wild-type. Analysis of the phenotype of PA14 RSCV indicated that these variants exhibited increased aggregation and attachment to glass tubes and polyvinylchloride plastic (PVC) as a result of enhanced surface
- 25 hydrophobicity. Consistent with these observations, several PA14 RSCV clones were hyperpilated when analysed by transmission electron microscopy. Moreover, examination of biofilms cultivated in flow chamber cells showed that PA14 RSCV formed more biofilm and faster than the wild-type strain. The biofilm formed by PA14 RSCV also showed increased resistance to tobramycin relative to wild-type PA14
- 30 biofilm. Similar results were obtained for several CF isolates using different antibiotics

(including tobramycin), suggesting that nonspecific antibiotic resistance acquired through phenotypic variation is a common mechanism in *P. aeruginosa*. Moreover, analysis of sputum samples taken from CF patients revealed that antibiotic treatment selects for antibiotic resistant variants. The frequency with which antibiotic resistant variants appeared was also affected by environmental stimuli. Environmental stimuli such as salt concentration, temperature, and bacterial media altered the frequency of appearance of resistant variants.

To identify components involved in the regulation of antibiotic resistance mediated by phenotypic variation, a library of PA14 chromosomal DNA was transferred into PA14 RSCV and screened for colonies displaying wild-type colony size and morphology. This led to the identification of a clone, pED202, that restored the colony, the autoagglutination, and attachment phenotypes of PA14 RSCV variants to wild-type. pED202 contained a single gene (designated *pvrR* for phenotype variant regulator) that showed sequence similarities to response regulator elements of the two-component regulatory system found in *Vibrio cholerae* response regulator VieA, and in *P. aeruginosa* strain PAO1 (ORF PA3947).

Consistent with the putative role of PvrR in the regulation of phenotypic switching, overexpression of PvrR from pED202 in wild-type PA14 resulted in reduced attachment to PVC plastic. Moreover, examination of the frequency of resistant variants obtained from kanamycin plates showed a reduction in the number of colonies resistant to antibiotic obtained from the PvrR overexpressing strain. An in-frame deletion of *pvrR* ($\Delta pvrR$) constructed in PA14 increased frequency of appearance of resistant variants on kanamycin plates with respect to the wild-type, confirming the involvement of *pvrR* in the regulation of phenotypic switching. These results suggested that PvrR might be acting upstream of the switch, since inactivation of *pvrR* by mutation did not result in conversion to the variant type.

Below we describe the cloning and characterization of PvrR, a regulator of biofilm-mediated antibiotic resistance and a target for compounds useful in antibacterial therapy, along with antibiotics, for the treatment of chronic infections and biofilm control in medical and industrial settings. In addition, we describe the identification of

open reading frame regions, designated ORF1 and ORF3, that flank the *pvrR* gene. The following examples are for the purposes of illustrating the invention, and should not be construed as limiting.

5 Appearance of Rough Small Colony Variants with Increased Antibiotic Resistance

When cultured under high concentrations of antibiotic, *Pseudomonas aeruginosa* PA14 was found to shift its development to a rough small colony phenotype, leading to the production of antibiotic resistant colonies. To induce such phenotypic variants, an overnight culture of *P. aeruginosa* strain PA14 (UCBPP-PA14) was inoculated onto

10 Luria-Bertani (LB) containing 200 µg/ml of kanamycin, incubated at 37°C for 48 hours, at which time, antibiotic resistant rough small variants were isolated. Antibiotic resistant colonies arose at a frequency of 10^{-6} - 10^{-7} . The colonies identified on these plates were one-tenth the size of wild type and exhibited a rough phenotype compared to the smooth colony type of wild-type PA14. One class of kanamycin resistant variants

15 (approximately 30%) exhibited a rough phenotype compared to the smooth colony type of wild-type PA14. When incubated for three to five days in LB media without antibiotic at room temperature, the rough phenotype reverted to the wild-type phenotype (Figure 1A), indicating that the phenotypic changes were transient, and not due to mutation. In addition to being resistant to kanamycin, (up to 40 times the susceptibility

20 level of the wild-type), 8 individual PA14 RSCV colonies tested were also resistant to amikacin (30 µg/ml), carbenicillin (300 µg/ml), gentamicin (30 µg/ml), tobramycin (10 µg/ml), and tetracycline (150 µg/ml). Consistent with this latter result, antibiotic resistant variants were also obtained at frequencies of about 10^{-7} by plating overnight cultures of PA14 on media containing similar concentrations of the antibiotics

25 mentioned above. Although RSCV colonies were smaller than wild-type, their small colony size was not a consequence of slow growth since the generation time of RSCV in liquid medium was not significantly different from that of the wild-type, even in LB containing 200 µg/ml kanamycin.

Phenotypic Changes Associated With Appearance of Resistance

To establish a connection between the phenotypic switch from wild-type to small variant colony and the emergence of antibiotic resistance, comparative attachment, agglutination, and biofilm formation studies of wild-type PA14 and PA14 RSCV were
5 conducted.

The results of these experiments showed that PA14 RSCV formed visible bacterial aggregates when overnight liquid cultures were left without shaking at room temperature (Figure 1B). Moreover, abundant bacterial aggregates formed when liquid cultures were grown with gentle agitation, indicating that PA14 RSCV had increased
10 cell-cell attachment compared to the wild-type phenotype.

In addition to the autoagglutination phenotype, PA14 RSCV developed a visible biofilm on the walls of glass tubes after overnight incubation in liquid culture. Wild-type PA14 failed to form a similar biofilm under these conditions. These results indicated that cell-surface interactions, as well as cell-cell interactions were increased in
15 the variant. Consistent with this observation, PA14 RSCV were found to have increased attachment to PVC plastic (Figure 1C) in assays conducted in 96-well microtiter plates. When reversion was induced in PA14 RSCV, the reverted bacteria showed wild-type levels of both agglutination and attachment to glass and PVC plastic.

To quantitatively assess differences between the strains, standard bacterial
20 attachment assays were performed in 96-well polyvinylchloride (PVC) plastic plates according to the methods described by O'Toole et al. (*Mol. Microbiol.* 30: 295, 1998). Overnight cultures of PA14 and PA14 RSCV were diluted to an OD₆₀₀ of 0.1 in fresh minimal M63 salts supplemented with glucose (0.3%), MgSO₄ (1 mM), and casamino acids (CAA, 0.5%). Aliquots of 100 µl were next dispensed into the wells of PVC
25 plastic microtiter plates and incubated for 6 hours at 37°C. The attachment of bacteria to the walls of the microtiter well was then detected by staining with 1% crystal violet dissolved in water. Dye not associated with bacteria was removed by thorough rinsing with water. Bacteria-associated dye was solubilized using 95% ethanol and absorbance was determined at OD₅₅₀.

In addition, since the ability of bacteria to attach to each other and to surfaces depends in part on the interaction of hydrophobic domains (Drumm et al., *J. Clin. Invest.* 84:1588, 1989), the hydrophobic surface properties of the wild-type and PA14 RSCV were determined using a standard salt aggregation test (Sherman et al., *Infect. Immun.* 5 49:797, 1985). 5×10^8 bacteria per ml in 0.025 ml were mixed on a microscope slide with an equal volume of ammonium sulfate in 0.002 M sodium phosphate, pH 6.8. The ammonium sulfate concentrations varied from 0.0625 M to 4.0 M, and the presence of salt-induced bacterial aggregation was monitored for 2 minutes at room temperature by phase-contrast microscopy. Agglutination in salt concentrations of less than 0.1 M is 10 taken as an indication of the presence of a hydrophobic bacterial surface. Hydrophilic surfaces were demonstrated by the agglutination of bacteria only in high salt concentrations (2.0 to 4.0 M).

The data obtained from the salt aggregation tests showed that PA14 RSCV were agglutinated at a lower salt concentration (0.125 M) compared to the wild-type PA14 15 (0.5 M), suggesting that PA14 RSCV has a higher degree of surface hydrophobicity than the wild-type. Therefore, the data indicated that a change in the hydrophobic properties of the surface of the bacteria was partially responsible for the general increase in surface attachment of the PA14 RSCV phenotypic variant. To further demonstrate the role of hydrophobicity in surface attachment, PA14 RSCV were cultured in the presence of 20 tetramethyl urea (TMU), a hydrophobic bond-breaking agent, at a concentration of 200 mM. Addition of TMU to the culture media was found to reduce the attachment of the phenotypic variant PA14 RSCV to wild-type levels, confirming the hydrophobic nature of the bacterial surface. TMU, at the concentration used in these assays, did not affect cell viability.

25 Transmission electron microscopic analysis of several PA14 RSCV clones revealed that they were hyperpiliated, which is consistent with the increased hydrophobicity and agglutination phenotypes. However, the various phenotypes of PA14 RSCV were not simply a consequence of hyperpiliation since a hyperpiliated mutant of *P. aeruginosa* PA14, *pilU*, exhibited only marginally enhanced 30 hydrophobicity and attachment to PVC plastic and did not exhibit enhanced resistance to

antibiotics (data not shown). These results are consistent with previous reports which indicated that phenotypic variation in Gram-negative bacteria involve changes in expression of a number of surface structures, outer membrane proteins, and lipopolysaccharides resulting in altered aggregation and colony morphology. Several
5 PA14 RSCV clones were tested in the experiments described above and all exhibited similar phenotypes. A single PA14 RSCV clone was therefore chosen for further analysis.

To determine whether the antibiotic resistant phenotype of PA14 RSCV is associated with altered biofilm formation, PA14 RSCV was cultured under biofilm-
10 forming conditions as follows. For biofilm characterization, PA14 RSCV biofilms were cultivated under continuous culture conditions in flow-chambers with channel dimensions of 12 by 52 by 2 mm. Flow media consisted of M63 supplemented with 0.5% casamino acids and 0.3% glucose. For measurement of biofilm resistance, bacteria were cultivated in flow-chambers with channel dimensions of 1 by 40 by 4 mm (Stovall
15 Inc., Greensboro, NC). In this case, flow media consisted of FAB medium (0.1 mM CaCl_2 , 0.01 mM Fe-EDTA, 0.15 mM NH_4SO_4 , 0.33 mM Na_2HPO_4 , 0.2 mM KH_2PO_4 and 1 mM MgCl_2) supplemented with casamino acids (0.5%) and sodium citrate (10 mM). Flow-cells in both cases were inoculated with 100-fold dilutions of overnight cultures of PA14 and PA14 RSCV carrying the green fluorescent protein (GFP) in
20 plasmid SMC21, a derivative of pSMC2 (Bloemberg et al., *Appl. Environ. Microbiol.* 63: 4543-4551, 1997). After inoculation, the medium flow was stopped for 1 hour. Medium flow was then resumed at a rate of 0.2 ml/min using a peristaltic pump (IsmaTec, Zurich, Switzerland), and the flow-cell system was incubated at 37° C. Analysis of biofilm spatial structures was performed using confocal scanning laser
25 microscopy (CSLM) using a Leica TCS SP system (Leica Lasertechnik, GmG, Heidelberg, Germany). Image analysis of antibiotic-treated biofilms was done in structures contained within serial section stacks of images delimited by freehand drawing. Pixel intensities unique to GFP-labeled bacteria and surrounding biofilm were established by the threshold limit technique. The volume (in μm^3) of individual biofilm

structures was determined from serial sections using ImageSpace software (Molecular Dynamics, Sunnyvale, CA).

The results from these studies showed that the PA14 RSCV phenotypic variant formed not only more biofilm than the wild-type strain, but also formed biofilm faster (RSCV microcolonies appeared 4-5 hours earlier than wild-type). Moreover, PA14 RSCV and wild-type PA14 displayed significantly different patterns of biofilm development. Wild-type PA14 initially formed regularly-spaced, flat, circular, microcolonies that eventually developed into ball-shaped microcolonies. In contrast, PA14 RSCV formed irregularly shaped three-dimensional structures that were densely packed with bacteria, without the typical microcolony morphology (Figure 1D). Finally, the biofilm structures formed by PA14 RSCV were larger in size than the wild-type microcolonies, and biofilms from PA14 RSCV contained more biomass than the wild-type.

To determine whether PA14 and PA14 RSCV biofilms exhibited antibiotic resistance that paralleled the resistance observed on plates containing antibiotic, established PA14 and PA14 RSCV biofilms grown in flow chambers were exposed to a continuous flow of tobramycin (200 µg/ml). Viable biomass was measured by CSLM analysis of GFP-tagged PA14 and PA14 RSCV cells using GFP expression as a viability marker as described previously (Figure 1E). Consistent with the results obtained in plates, the biofilm formed by PA14 RSCV was more resistant to tobramycin treatment than the wild-type PA14 biofilm.

Phenotypic variation is a common phenomenon in Gram-negative bacteria that often involves environmentally regulated changes in observable phenotypes produced by modifications in surface components. The effect that different environmental stimuli had on the appearance of kanamycin-resistant phenotypic variants was examined. Bacteria were grown in LB broth, or in supplemented LB with appropriate antibiotics at the indicated temperature with aeration. As shown in Figure 2A, a 40-fold increase in the frequency of appearance of resistant variants (not just PA14 RSCV) was observed on LB media supplemented with 85 mM NaCl as compared to the same medium without NaCl. Moreover, the frequency of variants increased 200-fold when plates were

incubated at 25°C compared to 37°C (Figure 2A). Finally, a dramatic 10^6 - fold increase was obtained on minimal M63 salts as compared to LB medium (Figure 2A). Minimal salt media consisted of M63 supplemented with 0.3% glucose, 1 mM MgSO_4 , and 0.5% casamino acids. Importantly, there was a correlation between the frequency of appearance of kanamycin resistant variants on plates and minimal inhibitory concentrations (MICs) of kanamycin in liquid culture for the wild-type PA14 using the culture conditions described above (Figure 2B). For example, the high frequency of resistant variants obtained on M63 correlated with the relatively high concentration of kanamycin (475 $\mu\text{g/ml}$) required to inhibit the growth of PA14 in M63 liquid medium (Figures 2A and 2B). These data indicated that the components involved in the formation of antibiotic resistant variants are differentially regulated by environmental signals. Moreover, the data indicated that the portion of the population that becomes resistant to antibiotics through phenotypic variation was largely dependent on environmental conditions.

Small Colony Variants in CF Sputum Samples

The presence of phenotypic variants with small colony phenotypes has been reported in cystic fibrosis (CF) patients (Haussler et al., *Clin. Infect. Dis.* 29:621, 1999). Emergence of this and other variant phenotypes in the CF lung has also been linked to prolonged antibiotic treatment (McNamara et al., *Int. J. Antimicrob. Agents* 14:117, 2000; Kahl et al., *J. Infect. Dis.* 177:1023, 1998). To investigate whether antibiotic treatment in *P. aeruginosa* CF infections results in selection for resistant variants, we looked for the presence of small colony variants in CF sputum samples.

Five CF sputum samples from the Clinical Microbiology Laboratory at Massachusetts General Hospital were suspended in 5 ml of 10 mM MgSO_4 . Serial dilutions of the samples were then plated onto ceftrimide agar plates with and without antibiotics. The plates were screened for the presence of *P. aeruginosa* after 24 and 48 hours of incubation at 37°C. The identity of the colonies was later confirmed by probing colony lifts with the exotoxin A gene from *P. aeruginosa*. To this end, the *EcoRI*-*HindIII* fragment of plasmid pRGI containing the *exoA* gene (Samadpour et al., *J. Clin.*

Microbiol. 26:2319-23, 1988) was gel isolated and labeled using a random priming kit (Boehringer, Mannheim, Indianapolis, Ind.). Colonies were transferred to nylon membranes and hybridizations were performed according to the manufacturer's recommendations (NEN Research Products, Boston, MA). Identification of colonies carrying the *exoA* gene was then performed using a Phosphorimager (Amersham Pharmacia Biotech Inc., Piscataway, NJ).

Five sputum samples obtained from five CF patients were evaluated for the presence of small colony variant bacteria. Two out of five sputum samples obtained from CF patients (patients 5 and 38) contained 100% rough small colony variants (Table 1) that reverted to a wild-type colony morphology upon prolonged incubation on antibiotic-free medium (Figure 3A). Importantly, both samples 5 and 38 corresponded to patients that were undergoing antibiotic treatment at the time the samples were obtained (intravenous (IV) amikacin/ceftazidime for two days and oral (O) levofloxacin/inhaled (I) tobramycin for six weeks respectively Table 1).

TABLE 1

	Sample 5	Sample 38	Sample 41	Sample 42	Sample 43
Antibiotic treatment of CF patients	Amikacin(IV) Ceftazidime(IV)	Tobramycin (I) Levofloxacin(O)	none	none	none
Small Colony variants in sputum sample (%)	100	100	< 0.11	0.00	< 0.12
Variants resistant to amikacin (%)	100	100	15	5	0.2
Variants resistant to gentamicin (%)	100	100	10	6.6	0.5
Variants resistant to tetracycline (%)	30	32	0	0	Not done
Variants resistant to tobramycin (%)	50	100	0.10	0	0.5

Table 1 shows the presence of small colony *P. aeruginosa* variants in sputum samples from five CF patients. The presence of *P. aeruginosa* antibiotic resistant small colony variants was determined by plating CF sputum samples on ceftrimide agar with and without the indicated antibiotics.

5 Moreover, there was 29% enrichment in small colony variants in samples taken on two consecutive days from the patient that was undergoing intravenous antibiotic treatment.

As shown in Table 1, 30-100% of the small colony variants present in samples 5 and 38 were resistant to four different antibiotics (amikacin, gentamicin, tetracycline, and tobramycin) at concentrations equal to or higher than the minimal bactericidal
10 concentration (MBC) of their respective reverted colonies. The proportion of small colony variants present in the samples that showed resistance to amikacin, gentamicin, tetracycline, and tobramycin was analyzed by simultaneously plating the sputum samples in ceftrimide agar with and without antibiotics. The data obtained were
15 compared to MBCs of the reverted colonies for the antibiotics in which variants were obtained. *In vitro* susceptibility (MBC) to the different antibiotics used during the assays was determined by a standard tube dilution procedure described by Bailey and Scott (*Diagnostic Microbiology*, 313-329, 1974).

Although the other three CF sputum samples (41, 42 and 43) appeared to contain
20 either a small proportion or no detectable small colony variants when plated on antibiotic free media, they did contain a considerable number (0.5–15%) of antibiotic resistant variants (Table 1). This discrepancy was due to the fact that it took the small colony variants 36-40 hours to form visible colonies, at which time the fast growing wild-type bacteria present in the sputum samples had overgrown the antibiotic free
25 plates. Resistant variants with small colony phenotypes obtained from plating CF isolates 42 and 43 on media containing tobramycin (a front-line antibiotic used for the treatment of *P. aeruginosa* infections) exhibited increased attachment to PVC plastic (Figure 3B).

Identification of the Phenotypic Variation Regulator Gene

Phenotypic variation is a common mechanism in Gram-negative bacteria, and involves changes in observable phenotypes produced by modifications in surface components such as fimbriae, flagella, outer membrane proteins, and lipopolysaccharides. In the mushroom pathogen *P. tolaasii*, Greewal et al. (*J. Bacteriol.* 177:4658, 1995) identified a two-component regulatory element responsible for the phenotypic switch from smooth to rough phenotype that involved changes in colony morphology and motility. Since the phenotype displayed by PA14 RSCV was transient and involved alterations in surface properties, we hypothesized that a regulatory component was also responsible for the phenotypic switch observed in PA14.

To identify this component, a genomic library of strain PA14 constructed in the cosmid vector pJSR1 (Rahme et al., *Science* 268:1899, 1995) was mobilized in masse into PA14 RSCV by triparental mating using helper strain pRK2013 (Figurski et al., *Proc. Natl. Acad. Sci. USA* 76:1648, 1979). The resulting transconjugants were screened visually for colonies showing wild-type size and morphology (smooth colony phenotype). Individual transconjugants that showed wild-type characteristics were used to isolate the corresponding cosmids which were then reintroduced into PA14 RSCV to confirm the reversion of the phenotype. Moreover, cosmid DNA from the transconjugants was digested to completion with the restriction enzymes *EcoRI*, *PstI*, and *HindIII* and separated by electrophoresis on a 0.7% agarose gel.

A total of 2,500 transconjugants were screened for colonies displaying wild-type PA14 colony size and morphology. Two transconjugants that showed wild-type phenotypes were isolated, indicating that the inserts contained in the cosmids were able to induce reversion from small colony variant to wild-type phenotype. Two cosmid clones were isolated and reintroduced in PA14 RSCV to test for restoration of wild-type phenotype, and both clones were found to be capable of greatly enhancing the rate of PA14 RSCV reversion to the wild-type phenotype. Restriction digest profiles obtained with *EcoRI*, *PstI*, and *HindIII* restriction enzymes showed the presence of a cosmid with the same insert in both cases, which was designated pED20. Although the PA14 RSCV phenotype was normally very stable in liquid culture (i.e., no wild-type revertants

observed when an overnight culture was plated on LB agar), the majority of the cells in a PA14 RSCV culture carrying pED20 formed wild-type colonies after overnight incubation.

Cosmid pED20 was then subcloned into the pUCP19 plasmid vector using a PstI
5 restriction digest. The clones obtained after transformation in *E. coli* were used to isolate plasmid DNA that was subsequently introduced into PA14 RSCV by electroporation. The resulting clones were screened visually for colonies showing wild-type size and morphology. Subcloning of pED20 produced pED202, which contained a 3.5-kb fragment, that restored the colony phenotype of PA14 RSCV variant to wild-type.
10 Clone pED202 restored attachment phenotypes (Figure 4A), as well as the colony morphology and autoagglutination phenotypes of PA14 RSCV variants to wild-type. The vector alone did not have any effect on the phenotypes analyzed.

DNA sequencing and sequence analysis of the pED202 insert was then performed. The DNA fragments used for sequencing were PCR amplified initially using
15 primers M13 and M13 reverse from the pUCP19 plasmid. Primers were later synthesized based on the sequencing data obtained. Sequencing data were analyzed using the DNASTar software (DNASTAR Inc., Madison, WI) to predict the open reading frames present in the pED202 3.5 kb insert. Sequence information was also compared with the sequence databases at the National Center for Biotechnology Information as well as to the *P. aeruginosa* PAO1 sequence generated by the *P. aeruginosa* genome
20 project (Cystic Fibrosis Foundation and PathoGenesis Corporation).

Analysis of the sequencing data obtained from clone pED202 showed that the clone contained only one intact open reading frame. The nucleotide and predicted amino acid sequences of the ORF (designated *pvrR* for phenotype variant regulator)
25 contained in clone pED202 were compared to the GenBank databases, and showed sequence similarities to response regulator elements of the two-component regulatory system. The search revealed 30% identity and 45% similarity in a 376 amino acid overlap to the *Vibrio cholerae* response regulator VieA, which is induced during intestinal infection in mouse. In addition, the ORF on pED202 showed 29% identity and
30 45% similarity to a probable two-component response regulator identified in *P.*

aeruginosa strain PAO1 (ORF PA3947). Interestingly, the region of the PA14 genome containing *pvrR* is not present in the fully sequence *P. aeruginosa* strain PAO1.

A homology search against domain sequences in the ProDom database (ProDom web site; <http://prodes.Toulouse.inra.fr/prodom>) identified 4 regions with high-scoring segment pairs in PvrR (Figure 4B). All 4 domains are also present in VieA and the PA01 putative response regulator (Figure 4B). Moreover, these 4 domains exhibit high levels of amino acid sequence similarity (30%-60%; Figure 4B). Sequence analysis of the regions located upstream and downstream of *pvrR* revealed the presence of two additional ORFs (designated *ORF1* and *ORF3* respectively; Figure 4C) with sequence homology to two-component regulatory elements.

The protein encoded by ORF1 has homology to probable sensor/response regulator hybrids from *P. aeruginosa* (35% identity and 49% similarity to ORF. PA2824), to the sensor protein RcsC (capsular synthesis regulator component C) from *Salmonella enterica* subsp. *enterica* serovar Typhi (30% identity and 51% similarity) and to a two-component sensor regulator (PheN) that modulates phenotypic switching in *P. tolaasii*, (31% identity and 45% similarity). The protein encoded by ORF3 shows 42% identity and 60% similarity to the GacS sensor kinase from *P. fluorescens*, and 41% identity and 59% similarity to the two-component sensor regulator that modulates phenotypic switching in *P. tolaasii* (PheN).

Figure 5G shows a nucleic acid sequence (SEQ ID NO:7) including polynucleotides identified in the ORF1 region (SEQ ID NOS:3, and 8-18), *pvrR* (SEQ ID NO:1), polynucleotides identified in the ORF3 region (SEQ ID NOS:5, and 30-34), and the intergenic regions. The start and stop codons for each open reading frame are indicated by highlighting. Figures 5B and 6A-K show the nucleotide sequences of several open reading frames identified in the ORF1 region. The deduced amino acid sequence of these open reading frames are shown in Figures 5E (SEQ ID NO:4) and 6L-6V (SEQ ID NOS:19-29).

Additionally, Figure 5C shows the nucleic acid sequence (SEQ ID NO:5) of one of several open reading frames identified in the ORF3 region. The deduced amino acid sequence of the polypeptide encoded by this nucleotide sequence is shown in Figure 5F

(SEQ ID NO:6). Figures 7A-7E (SEQ ID NOS:30-34) show the nucleotide sequences of several additional open reading frames identified in the ORF3 region. The deduced amino acid sequence of the polypeptides encoded by these nucleotide sequences are shown in Figures 7F-7J.

5 To determine whether *pvrR* or a highly similar *pvrR* homolog was present in the other *P. aeruginosa* strains, PCR analysis of 14 *P. aeruginosa* strains was performed using *pvrR*-specific primers. The specificity of the PCR products obtained was subsequently confirmed by Southern blotting and hybridization with a *pvrR*-specific probe. Results showed that 7 out of 7 CF isolates, 2 out of 3 clinical isolates and 3 out
10 of 4 standard *P. aeruginosa* laboratory strains contained the *pvrR* gene fragment or a highly similar fragment (data not shown).

PvrR Overexpression

Consistent with the putative role of *PvrR* in the regulation of phenotypic
15 switching, overexpression of PvrR from pED202 resulted in a 6-fold reduction in the frequency of resistant variants obtained after plating overnight cultures on kanamycin (200 µg/ml) plates compared to wild-type (Figure 4D). Plasmid pED202, containing the *pvrR* gene was introduced into wild-type PA14 by electroporation using standard methods. Frequency of appearance of kanamycin resistant variants and attachment to
20 96-well PVC plates was assayed as described above. Interestingly, the PvrR overexpressing strain also caused a 2.5-fold reduction in attachment to PVC plastic with respect to the strain carrying the vector alone (Figure 4E).

pvrR Deletion Analysis

25 Since PvrR is involved in the regulation of the phenotypic switch from wild-type to phenotypic variant, a mutation in *pvrR* would be expected to alter the proportion of resistant variants present in the PA14 population. To test this hypothesis, a 914 bp in-frame deletion within *pvrR* (denoted “Δ*pvrR*”) was generated by replacing 2.33 kb of the wild-type sequence of the *pvrR* gene with a 1.416 kb fragment amplified by PCR.
30 The PCR-amplified DNA fragment was subcloned into the *Xba*I and *Sma*I restriction

sites of the positive selection suicide vector pCVD442 to generate pED167. Plasmid pED167 was then used in an allelic exchange procedure to introduce the fragment containing the deleted copy of *pvrR* into the homologous region of the PA14 chromosome, creating strain ED78. The deletion was confirmed by sequencing a PCR
5 fragment containing *pvrR*.

This deletion of *pvrR* ($\Delta pvrR$) in PA14 resulted in an increased frequency of appearance of resistant variants on kanamycin plates with respect to the wild-type (Figure 4F), confirming the involvement of *pvrR* in the regulation of phenotypic switching. The observation that 100% of the variants expressing wild-type *pvrR*
10 reverted to the wild-type phenotype implicates PvrR is inducing reversion from variant to wild-type phenotypes. These results indicated that PvrR may be acting upstream of the switch, since inactivation of *pvrR* by mutation was not found to result in conversion to the variant type.

15 Isolation of Additional Biofilm Regulator Genes

Based on the nucleotide and amino acid sequences described herein, the isolation and identification of additional coding sequences of genes regulating the formation of microbial biofilm is made possible using standard strategies and techniques that are well known in the art. For example, any microbe that possesses the ability to form a biofilm
20 can serve as the nucleic acid source for the molecular cloning of such a gene, and these sequences are identified as ones encoding a protein exhibiting structures, properties, or activities associated with biofilm formation, such as the PvrR (Figure 5D, SEQ ID NO:2), or any of the polynucleotides identified in the ORF1 (SEQ ID NOS:3 and 8-18) and ORF3 (SEQ ID NOS:5 and 30-34) regions.

In one particular example of such an isolation technique, any one of the nucleotide sequences described herein, including *pvrR* (Figure 5A, SEQ ID NO:1), *ORF1* (Figure 5B, SEQ ID NO:3), or *ORF3* (Figure 5C, SEQ ID NO:5) may be used, together with conventional methods of nucleic acid hybridization screening. Such hybridization techniques and screening procedures are well known to those skilled in the
30 art and are described, for example, in Benton and Davis (*Science* 196:180, 1977);

Grunstein and Hogness (*Proc. Natl. Acad. Sci., USA* 72:3961, 1975); Ausubel et al. (*Current Protocols in Molecular Biology*, Wiley Interscience, New York, 2001); Berger and Kimmel (*Guide to Molecular Cloning Techniques*, 1987, Academic Press, New York); and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York. In one particular example, all or part of the *pvrR*, *ORF1*, or *ORF3* sequences (described herein) may be used as a probe to screen a recombinant DNA library for genes having sequence identity to the *pvrR*, *ORF1*, or *ORF3* genes. Hybridizing sequences are detected by plaque or colony hybridization according to standard methods.

10 Alternatively, using all or a portion of the amino acid sequences of *PvrR*, *ORF1*, or *ORF3*, one may readily design *pvrR*, *ORF1*, or *ORF3* gene-specific oligonucleotide probes, including degenerate oligonucleotide probes (i.e., a mixture of all possible coding sequences for a given amino acid sequence). These oligonucleotides may be based upon the sequence of either DNA strand and any appropriate portion of the *pvrR*,
15 *ORF1*, or *ORF3* sequences. General methods for designing and preparing such probes are provided, for example, in Ausubel et al. (supra), and Berger and Kimmel, *Guide to Molecular Cloning Techniques*, 1987, Academic Press, New York. These oligonucleotides are useful for *pvrR*, *ORF1*, or *ORF3* gene isolation, either through their use as probes capable of hybridizing to *pvrR*, *ORF1*, or *ORF3* complementary sequences
20 or as primers for various amplification techniques, for example, polymerase chain reaction (PCR) cloning strategies. If desired, a combination of different, detectably-labelled oligonucleotide probes may be used for the screening of a recombinant DNA library. Such libraries are prepared according to methods well known in the art, for example, as described in Ausubel et al. (supra), or they may be obtained from
25 commercial sources.

As discussed above, sequence-specific oligonucleotides may also be used as primers in amplification cloning strategies, for example, using PCR. PCR methods are well known in the art and are described, for example, in *PCR Technology*, Erlich, ed., Stockton Press, London, 1989; *PCR Protocols: A Guide to Methods and Applications*,
30 Innis et al., eds., Academic Press, Inc., New York, 1990; and Ausubel et al. (supra).

Primers are optionally designed to allow cloning of the amplified product into a suitable vector, for example, by including appropriate restriction sites at the 5' and 3' ends of the amplified fragment (as described herein). If desired, nucleotide sequences may be isolated using the PCR "RACE" technique, or Rapid Amplification of cDNA Ends (see, 5 e.g., Innis et al. (supra)). By this method, oligonucleotide primers based on a desired sequence are oriented in the 3' and 5' directions and are used to generate overlapping PCR fragments. These overlapping 3'- and 5'-end RACE products are combined to produce an intact full-length cDNA. This method is described in Innis et al. (supra); and Frohman et al., *Proc. Natl. Acad. Sci. USA* 85:8998, 1988.

10 Partial sequences, e.g., sequence tags, are also useful as hybridization probes for identifying full-length sequences, as well as for screening databases for identifying previously unidentified related virulence genes.

In general, the invention includes any nucleic acid sequence which may be isolated as described herein or which is readily isolated by homology screening or PCR 15 amplification using any of the nucleic acid sequences disclosed herein such as those shown in Figures 5A, 5C, 5G, 6A-K, or 7A-7E.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding PvrR, ORF1, or ORF3, some bearing minimal similarity to the polynucleotide sequences of any known 20 and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally-occurring *pvrR*, *ORF1*, or *ORF3*, and all such variations are to be 25 considered as being specifically disclosed.

Although nucleotide sequences which encode PvrR, ORF1, ORF3, or their variants are preferably capable of hybridizing to the nucleotide sequence of the naturally-occurring *pvrR*, *ORF1*, or *ORF3* under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding PvrR, 30 ORF1, ORF3, or their derivatives possessing a substantially different codon usage, e.g.,

inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding PvrR, ORF1, ORF3, and their derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode PvrR, ORF1, ORF3, or fragments thereof generated entirely by synthetic chemistry.

After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding any one of PvrR, ORF1, ORF3, or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in Figure 5A, 5B, 5C, 5G, 6A-6K, or 7A-7E and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G. M. and S. L. Berger (1987) *Methods Enzymol.* 152:399; Kimmel, A. R. (1987) *Methods Enzymol.* 152:507) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30 °C, more preferably of at least about 37 °C, and most preferably of at least about 42 °C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a

preferred embodiment, hybridization will occur at 30 °C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37 °C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 µg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42 °C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50% formamide, and 200 µg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25 °C, more preferably of at least about 42 °C, and most preferably of at least about 68 °C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F. M. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York N.Y., unit 7.7)

Polypeptide Expression

In general, polypeptides of the invention (e.g., PvrR, ORF1, or ORF3 as shown in Figures 5D, 5E, 5F, 6L-6V, or 7F-7J) may be produced by transformation of a

suitable host cell with all or part of a polypeptide-encoding nucleic acid molecule or fragment thereof in a suitable expression vehicle.

Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems may be used to provide the recombinant protein. The precise host cell used is not critical to the invention. A polypeptide of the invention may be produced in a prokaryotic host (e.g., *E. coli*) or in a eukaryotic host (e.g., *Saccharomyces cerevisiae*, insect cells, e.g., Sf21 cells, or mammalian cells, e.g., NIH 3T3, HeLa, or preferably COS cells). Such cells are available from a wide range of sources (e.g., the American Type Culture Collection, Rockland, MD; also, see, e.g., Ausubel et al., supra). The method of transformation or transfection and the choice of expression vehicle will depend on the host system selected. Transformation and transfection methods are described, e.g., in Ausubel et al. (supra); expression vehicles may be chosen from those provided, e.g., in Cloning Vectors: A Laboratory Manual (P.H. Pouwels et al., 1985, Supp. 1987).

One particular bacterial expression system for polypeptide production is the *E. coli* pET expression system (Novagen, Inc., Madison, WI). According to this expression system, DNA encoding a polypeptide is inserted into a pET vector in an orientation designed to allow expression. Since the gene encoding such a polypeptide is under the control of the T7 regulatory signals, expression of the polypeptide is achieved by inducing the expression of T7 RNA polymerase in the host cell. This is typically achieved using host strains which express T7 RNA polymerase in response to IPTG induction. Once produced, recombinant polypeptide is then isolated according to standard methods known in the art, for example, those described herein.

Another bacterial expression system for polypeptide production is the pGEX expression system (Pharmacia). This system employs a GST gene fusion system which is designed for high-level expression of genes or gene fragments as fusion proteins with rapid purification and recovery of functional gene products. The protein of interest is fused to the carboxyl terminus of the glutathione S-transferase protein from *Schistosoma japonicum* and is readily purified from bacterial lysates by affinity chromatography using Glutathione Sepharose 4B. Fusion proteins can be recovered under mild

conditions by elution with glutathione. Cleavage of the glutathione S-transferase domain from the fusion protein is facilitated by the presence of recognition sites for site-specific proteases upstream of this domain. For example, proteins expressed in pGEX-2T plasmids may be cleaved with thrombin; those expressed in pGEX-3X may be
5 cleaved with factor Xa.

Once the recombinant polypeptide of the invention is expressed, it is isolated, e.g., using affinity chromatography. In one example, an antibody (e.g., produced as described herein) raised against a polypeptide of the invention may be attached to a column and used to isolate the recombinant polypeptide. Lysis and fractionation of
10 polypeptide-harboring cells prior to affinity chromatography may be performed by standard methods (see, e.g., Ausubel et al., supra).

Once isolated, the recombinant protein can, if desired, be further purified, e.g., by high performance liquid chromatography (see, e.g., Fisher, *Laboratory Techniques In Biochemistry And Molecular Biology*, eds., Work and Burdon, Elsevier, 1980).

15 Polypeptides of the invention, particularly short peptide fragments, can also be produced by chemical synthesis (e.g., by the methods described in *Solid Phase Peptide Synthesis*, 2nd ed., 1984 The Pierce Chemical Co., Rockford, IL). Also included in the invention are polypeptides which are modified in ways which do not abolish their pathogenic activity (assayed, for example as described herein). Such changes may
20 include certain mutations, deletions, insertions, or post-translational modifications, or may involve the inclusion of any of the polypeptides of the invention as one component of a larger fusion protein.

The invention further includes analogs of any naturally-occurring polypeptide of the invention. Analogs can differ from the naturally-occurring the polypeptide of the
25 invention by amino acid sequence differences, by post-translational modifications, or by both. Analogs of the invention will generally exhibit at least 85%, more preferably 90%, and most preferably 95% or even 99% identity with all or part of a naturally-occurring amino acid sequence of the invention. The length of sequence comparison is at least 15 amino acid residues, preferably at least 25 amino acid residues, and more preferably
30 more than 35 amino acid residues. Again, in an exemplary approach to determining the

degree of identity, a BLAST program may be used, with a probability score between e^{-3} and e^{-100} indicating a closely related sequence. Modifications include in vivo and in vitro chemical derivatization of polypeptides, e.g., acetylation, carboxylation, phosphorylation, or glycosylation; such modifications may occur during polypeptide synthesis or processing or following treatment with isolated modifying enzymes. 5
Analogues can also differ from the naturally-occurring polypeptides of the invention by alterations in primary sequence. These include genetic variants, both natural and induced (for example, resulting from random mutagenesis by irradiation or exposure to ethanemethylsulfate or by site-specific mutagenesis as described in Sambrook, Fritsch 10 and Maniatis, *Molecular Cloning: A Laboratory Manual* (2d ed.), CSH Press, 1989, or Ausubel et al., supra). Also included are cyclized peptides, molecules, and analogs which contain residues other than L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g., β or γ amino acids.

In addition to full-length polypeptides, the invention also includes fragments of any one of the polypeptides of the invention. As used herein, the term "fragment," 15 means at least 5, preferably at least 20 contiguous amino acids, preferably at least 30 contiguous amino acids, more preferably at least 50 contiguous amino acids, and most preferably at least 60 to 80 or more contiguous amino acids. Fragments of the invention can be generated by methods known to those skilled in the art or may result from normal 20 protein processing (e.g., removal of amino acids from the nascent polypeptide that are not required for biological activity or removal of amino acids by alternative mRNA splicing or alternative protein processing events). The aforementioned general techniques of polypeptide expression and purification can also be used to produce and isolate useful peptide fragments or analogs (described herein).

25

Antibodies

The polypeptides disclosed herein or variants thereof or cells expressing them can be used as an immunogen to produce antibodies immunospecific for such polypeptides. "Antibodies" as used herein include monoclonal and polyclonal 30 antibodies, chimeric, single chain, simianized antibodies and humanized antibodies, as

well as Fab fragments, including the products of an Fab immunoglobulin expression library.

To generate antibodies, a coding sequence for a polypeptide of the invention may be expressed as a C-terminal fusion with glutathione S-transferase (GST) (Smith et al.,
5 *Gene* 67:31, 1988). The fusion protein is purified on glutathione-Sepharose beads, eluted with glutathione, cleaved with thrombin (at the engineered cleavage site), and purified to the degree necessary for immunization of rabbits. Primary immunizations are carried out with Freund's complete adjuvant and subsequent immunizations with Freund's incomplete adjuvant. Antibody titres are monitored by Western blot and
10 immunoprecipitation analyses using the thrombin-cleaved protein fragment of the GST fusion protein. Immune sera are affinity purified using CNBr-Sepharose-coupled protein. Antiserum specificity is determined using a panel of unrelated GST proteins.

As an alternate or adjunct immunogen to GST fusion proteins, peptides corresponding to relatively unique immunogenic regions of a polypeptide of the
15 invention may be generated and coupled to keyhole limpet hemocyanin (KLH) through an introduced C-terminal lysine. Antiserum to each of these peptides is similarly affinity purified on peptides conjugated to BSA, and specificity tested in ELISA and Western blots using peptide conjugates, and by Western blot and immunoprecipitation using the polypeptide expressed as a GST fusion protein.

20 Alternatively, monoclonal antibodies which specifically bind any one of the polypeptides of the invention are prepared according to standard hybridoma technology (see, e.g., Kohler et al., *Nature* 256:495, 1975; Kohler et al., *Eur. J. Immunol.* 6:511, 1976; Kohler et al., *Eur. J. Immunol.* 6:292, 1976; Hammerling et al., In *Monoclonal Antibodies and T Cell Hybridomas*, Elsevier, NY, 1981; Ausubel et al., *supra*). Once
25 produced, monoclonal antibodies are also tested for specific recognition by Western blot or immunoprecipitation analysis (by the methods described in Ausubel et al., *supra*). Antibodies which specifically recognize the polypeptide of the invention are considered to be useful in the invention; such antibodies may be used, e.g., in an immunoassay. Alternatively monoclonal antibodies may be prepared using the polypeptide of the

invention described above and a phage display library (Vaughan et al., *Nature Biotech* 14:309, 1996).

Preferably, antibodies of the invention are produced using fragments of the polypeptides disclosed herein which lie outside generally conserved regions and appear likely to be antigenic, by criteria such as high frequency of charged residues. In one specific example, such fragments are generated by standard techniques of PCR and cloned into the pGEX expression vector (Ausubel et al., *supra*). Fusion proteins are expressed in *E. coli* and purified using a glutathione agarose affinity matrix as described in Ausubel et al. (*supra*). To attempt to minimize the potential problems of low affinity or specificity of antisera, two or three such fusions are generated for each protein, and each fusion is injected into at least two rabbits. Antisera are raised by injections in a series, preferably including at least three booster injections.

Antibodies against any of the polypeptides described herein may be employed to treat bacterial infections, for example, those infections involving biofilm formation. Thus, among others, antibodies against, for example, polypeptides of PvrR (SEQ ID NO: 2), ORF1 (SEQ ID NO: 4), or ORF3 (SEQ ID NO: 6) shown respectively in Figures 5D, E, or F may be employed to treat infections, particularly bacterial infections and especially chronic infections associated with CF or biofilm formation associated with indwelling medical devices, conjunctivitis, pneumonia, and bacteremia.

Diagnostics

In another embodiment, antibodies which specifically bind any of the polypeptides described herein may be used for the diagnosis of bacterial infection. A variety of protocols for measuring such polypeptides, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing bacterial infections.

In another aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding *pvrR*, *ORF1*, *ORF3*, or closely related molecules may be used to identify nucleic acid sequences which encode its gene product. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a

conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding PvrR, ORF1, or ORF3 allelic variants, or related sequences.

5 In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a
10 disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents. Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan et al., U.S. Pat. No. 5,474,796; Schena et al., *Proc. Natl. Acad. Sci.* 93:10614, 1996; Baldeschweiler et al., PCT application WO95/251116, 1995; Shalon, D. et al., PCT application WO95/35505, 1995; Heller et al., *Proc. Natl. Acad. Sci.* 94:2150, 1997; and Heller et al., U.S. Pat. No. 5,605,662.)
15

Screening Assays

As discussed above, we have identified a biofilm regulator gene, *pvrR*, of *P. aeruginosa* that mediates biofilm formation and antibiotic resistance by a microbe.

20 Based on this discovery, we have developed screening assays for identifying compounds that enhance or inhibit the action of a polypeptide or the expression of a nucleic acid sequence of the invention. The method of screening may involve high-throughput techniques.

Any number of methods are available for carrying out such screening assays. In
25 one working example, candidate compounds are added at varying concentrations to the culture medium of pathogenic cells expressing one of the nucleic acid sequences of the invention. Gene expression is then measured, for example, by standard Northern blot analysis (Ausubel et al., *supra*) or RT-PCR, using any appropriate fragment prepared from the nucleic acid molecule as a hybridization probe. The level of gene expression in
30 the presence of the candidate compound is compared to the level measured in a control

culture medium lacking the candidate molecule. A compound which promotes an increase in the expression of the *pvrR* gene or functional equivalent is considered useful in the invention; such a molecule may be used, for example, as a therapeutic to combat the pathogenicity of an infectious organism, for example, by decreasing its ability to
5 form a biofilm and rendering it susceptible to antibiotic treatment.

In another working example, the effect of candidate compounds may be measured at the level of polypeptide production using the same general approach and standard immunological techniques, such as Western blotting or immunoprecipitation with an antibody specific for a biofilm regulator polypeptide, such as PvrR. For
10 example, immunoassays may be used to detect or monitor the expression of at least one of the polypeptides of the invention in a microbial organism. Polyclonal or monoclonal antibodies (produced as described above) which are capable of binding to such a polypeptide may be used in any standard immunoassay format (e.g., ELISA, Western blot, or RIA assay) to measure the level of the polypeptide. A compound which
15 promotes an increase in the expression of the polypeptide is considered particularly useful. Again, such a molecule may be used, for example, as a therapeutic to combat the biofilm formation of an organism as is described above.

In yet another working example, candidate compounds may be screened for those which specifically bind to and agonize a PvrR polypeptide (a polypeptide having
20 the amino acid sequences shown in Figure 5D) of the invention. The efficacy of such a candidate compound is dependent upon its ability to interact with the PvrR polypeptide or functional equivalent thereof. Such an interaction can be readily assayed using any number of standard binding techniques and functional assays (e.g., those described in Ausubel et al., supra). For example, a candidate compound may be tested *in vitro* for
25 interaction and binding with a polypeptide of the invention and its ability to modulate biofilm formation may be assayed by any standard assay (e.g., those described herein).

In one particular working example, a candidate compound that binds to a polypeptide (e.g., PvrR) may be identified using a chromatography-based technique. For example, a recombinant polypeptide of the invention may be purified by standard
30 techniques from cells engineered to express the polypeptide (e.g., those described above)

and may be immobilized on a column. A solution of candidate compounds is then passed through the column, and a compound specific for the pathogenicity polypeptide (e.g, biofilm regulator polypeptide) is identified on the basis of its ability to bind to the pathogenicity polypeptide (e.g, biofilm regulator polypeptide) and be immobilized on the column. To isolate the compound, the column is washed to remove non-specifically bound molecules, and the compound of interest is then released from the column and collected. Compounds isolated by this method (or any other appropriate method) may, if desired, be further purified (e.g., by high performance liquid chromatography). In addition, these candidate compounds may be tested for their ability to render a pathogen incapable of forming a biofilm (e.g., as described herein). Compounds isolated by this approach may also be used, for example, as therapeutics to treat or prevent the onset of a pathogenic infection, disease, or both. Compounds which are identified as binding to pathogenicity polypeptides (e.g, biofilm regulator polypeptides) with an affinity constant less than or equal to 10 mM are considered particularly useful in the invention.

Potential agonists include organic molecules, peptides, peptide mimetics, polypeptides, and antibodies that bind to a nucleic acid sequence or polypeptide of the invention (e.g, biofilm regulator polypeptides) and thereby increase its activity. Potential agonists also include small molecules that bind to and occupy the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented.

Compounds that decrease only antibiotic resistance of a microbe are also identified by monitoring reversion of bacterial colonies from the antibiotic resistant phenotype to the wild-type susceptible phenotype. In one working example, screens for compounds that increase reversion rate are conducted by plating antibiotic resistant variant bacteria on antibiotic-free media in the presence or absence of a candidate compound. The plates are then cultured using standard methods. The plates are then visually inspected for revertants, i.e., bacterial colonies having a wild-type phenotype. The number of wild-type phenotype bacterial colonies is compared between plates cultured in the presence or absence of a candidate compound. Compounds that increase

the number of wild-type revertants, relative to the number of wild-type revertants on a control plate without the compound, are taken as useful in the invention.

Additionally, compounds that decrease antibiotic resistance are identified by monitoring for an increase in the susceptibility of bacteria to antibiotics. In yet another working example, compounds that decrease antibiotic resistance are identified by plating wild-type bacteria on antibiotic containing plates in the presence or absence of a candidate compound. The plates are cultured using standard methods, and then visually inspected for bacterial colonies. The number of antibiotic resistant bacterial colonies is compared between plates cultured in the presence or absence of a candidate compound.

Compounds that decrease the number of antibiotic resistant variant colonies, relative to the number of antibiotic resistant variant colonies on a control plate without the compound, are taken as useful in the invention.

In another working example, a gene that regulates biofilm formation is identified by monitoring its activity or activity of its encoded polypeptide, when mutated. Bacteria are mutagenized using standard methods, such as transposon mutagenesis. Mutagenized and wild-type bacteria are then plated on antibiotic containing plates. These plates are cultured using standard methods, and then are visually inspected for the presence of antibiotic resistant variant bacteria. The number of antibiotic resistant variant bacterial colonies (e.g., small colony variants) is compared between mutagenized bacterial plates and wild-type control plates. This comparison is typically conducted when variant colonies begin to appear on the wild-type plate. A decrease or increase in the number of antibiotic resistant variant bacterial colonies (e.g., small colony variants) on a plate containing mutagenized bacteria is taken as an indication of the presence of a genetic mutation in a gene that regulates biofilm formation. The mutated gene is then identified according to standard methods.

In yet another working example, a gene that regulates biofilm or phenotype-mediated antibiotic resistance is identified as follows. For example, a candidate gene (e.g., as part of a genomic library) is introduced into a variant host cell (e.g., *Pseudomonas aeruginosa* PA14 RSCV). Next, the transformed host cell is monitored for reversion from the rough small colony variant phenotype to wild-type. The plates

are then cultured using standard methods and monitored for the appearance of colonies with a wild-type phenotype. The number of wild-type phenotype bacterial colonies is then compared between plates containing transformants and variants carrying the vector alone. An increase in the number of wild-type revertants, relative to the number of wild-type revertants on a control plate with the vector alone, identifies a gene that regulates biofilm formation or phenotype-mediated antibiotic resistance. A gene identified using this method is subsequently isolated using standard procedures known in the art.

In another working example, small colony phenotypic variants are plated on an appropriate antibiotic medium (for example, those described herein) in the presence of a candidate compound and reversion to wild-type is monitored. Compounds that promote reversion from PA14 RSCV to wild-type are taken as being useful in the invention.

In another working example, a gene that regulates or is involved in phenotype-mediated or biofilm-mediated antibiotic resistance or biofilm formation is identified as follows. Bacteria are mutagenized using standard methods, such as transposon mutagenesis. Mutagenized bacteria are then plated on Trypticase Soy Agar (TSA) plates containing antibiotic. These plates are cultured using standard methods, and then inspected for bacterial growth. A decrease in the number of bacterial colonies or their absence on a mutagenized plate, relative to a control plate containing non-mutagenized bacteria identifies the presence of a genetic mutation in a gene that regulates phenotype-mediated or biofilm-mediated antibiotic resistance and biofilm formation. A gene identified using this method is subsequently isolated using standard procedures known in the art.

In another working example, a gene that regulates or is involved in phenotype-mediated or biofilm-mediated antibiotic resistance or biofilm formation is identified as follows. Bacteria are mutagenized using standard methods, such as transposon mutagenesis. Mutagenized bacteria are then transferred to Trypticase Soy Broth (TSB) liquid culture media containing an antibiotic. The bacteria are then cultured using standard methods, and the cultures are inspected for the presence of bacterial growth. Bacterial growth is compared between mutagenized cultures and wild-type control cultures. Bacterial growth can be identified, for example, by visual inspection, by

measuring optical density at 600 nm, or by other standard methods. The inability of a mutant to grow in liquid culture with antibiotics indicates the presence of a genetic mutation in a gene that regulates or is involved in phenotype-mediated or biofilm-mediated antibiotic resistance and biofilm formation. A gene identified using this method is subsequently isolated using standard procedures known in the art.

In another working example, a gene that regulates or is involved in phenotype-mediated or biofilm-mediated antibiotic resistance or biofilm formation is identified as follows. Bacteria are mutagenized using standard methods, such as transposon mutagenesis. Mutagenized bacteria are then plated on TSA plates containing antibiotic. These plates are cultured using standard methods, and then inspected for bacterial growth. The inability of a mutant to grow in TSA supplemented with antibiotics is taken as an indication of the presence of a genetic mutation in a gene that regulates or is involved in phenotype-mediated or biofilm-mediated resistance and biofilm formation. A gene identified using this method is subsequently isolated using standard procedures known in the art.

In another working example, a gene that regulates or is involved in phenotype-mediated or biofilm-mediated antibiotic resistance or biofilm formation is identified as follows. Bacteria are mutagenized using standard methods, such as transposon mutagenesis. Mutagenized bacteria are then transferred to liquid culture media TSB containing an antibiotic. The bacteria are then cultured using standard methods, and the cultures are inspected for the presence of bacterial growth. Bacterial growth is compared between mutagenized cultures and wild-type control cultures. Bacterial growth can be identified, for example, by visual inspection, by measuring optical density at 600 nm, or by other standard methods. The inability of a mutant to grow in liquid culture with antibiotics indicates the presence of a genetic mutation in a gene that regulates or is involved in phenotype-mediated or biofilm-mediated antibiotic resistance and biofilm formation. A gene identified using this method is subsequently isolated using standard procedures known in the art.

Each of the DNA sequences provided herein may also be used in the discovery and development of antipathogenic compounds (e.g., antibiotics). The encoded protein,

upon expression, can be used as a target for the screening of antibacterial drugs.

Additionally, the DNA sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The antagonists and agonists of the invention may be employed, for instance, to inhibit and treat a variety of bacterial infections, for example, those involving biofilm formation.

Optionally, compounds identified in any of the above-described assays may be confirmed as useful in conferring protection against the development of a pathogenic infection in any standard animal model (e.g., the mouse-burn assay described herein) and, if successful, may be used as anti-pathogen therapeutics (e.g, antibiotics).

Small molecules of the invention preferably have a molecular weight below 2,000 daltons, more preferably between 300 and 1,000 daltons, and most preferably between 400 and 700 daltons. It is preferred that these small molecules are organic molecules.

Test Compounds and Extracts

In general, compounds capable of reducing pathogenic virulence (e.g., reducing biofilm formation) are identified from large libraries of both natural product or synthetic (or semi-synthetic) extracts or chemical libraries according to methods known in the art. Those skilled in the field of drug discovery and development will understand that the precise source of test extracts or compounds is not critical to the screening procedure(s) of the invention. Accordingly, virtually any number of chemical extracts or compounds can be screened using the methods described herein. Examples of such extracts or compounds include, but are not limited to, plant-, fungal-, prokaryotic- or animal-based extracts, fermentation broths, and synthetic compounds, as well as modification of existing compounds. Numerous methods are also available for generating random or directed synthesis (e.g., semi-synthesis or total synthesis) of any number of chemical compounds, including, but not limited to, saccharide-, lipid-, peptide-, and nucleic acid-

based compounds. Synthetic compound libraries are commercially available from Brandon Associates (Merrimack, NH) and Aldrich Chemical (Milwaukee, WI).

Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are commercially available from a number of sources, including Biotics (Sussex, UK), Xenova (Slough, UK), Harbor Branch Oceanographics Institute (Ft. Pierce, FL), and PharmaMar, U.S.A. (Cambridge, MA). In addition, natural and synthetically produced libraries are produced, if desired, according to methods known in the art, e.g., by standard extraction and fractionation methods. Furthermore, if desired, any library or compound is readily modified using standard chemical, physical, or biochemical methods.

In addition, those skilled in the art of drug discovery and development readily understand that methods for dereplication (e.g., taxonomic dereplication, biological dereplication, and chemical dereplication, or any combination thereof) or the elimination of replicates or repeats of materials already known for their anti-pathogenic activity should be employed whenever possible.

When a crude extract is found to have an anti-pathogenic or anti-virulence activity, or a binding activity, further fractionation of the positive lead extract is necessary to isolate chemical constituents responsible for the observed effect. Thus, the goal of the extraction, fractionation, and purification process is the careful characterization and identification of a chemical entity within the crude extract having anti-pathogenic activity. Methods of fractionation and purification of such heterogeneous extracts are known in the art. If desired, compounds shown to be useful agents for the treatment of pathogenicity are chemically modified according to methods known in the art.

Pharmaceutical Therapeutics

The invention provides a simple means for identifying compounds (including peptides, small molecule inhibitors, and mimetics) capable of inhibiting the pathogenicity (e.g., biofilm formation) of a pathogen. Accordingly, a chemical entity discovered to have medicinal value using the methods described herein is useful as a

drug or as information for structural modification of existing anti-pathogenic compounds, e.g., by rational drug design. Such methods are useful for screening compounds having an effect on a variety of pathogens that form biofilms including, but not limited to, bacteria. Examples of pathogenic bacteria include, without limitation,

5 *Aerobacter, Aeromonas, Acinetobacter, Agrobacterium, Bacillus, Bacteroides, Bartonella, Bortella, Brucella, Calymmatobacterium, Campylobacter, Citrobacter, Clostridium, Cornyebacterium, Enterobacter, Enterococcus, Escherichia, Francisella, Haemophilus, Hafnia, Helicobacter, Klebsiella, Legionella, Listeria, Morganella, Moraxella, Proteus, Providencia, Pseudomonas, Salmonella, Serratia, Shigella,*

10 *Staphylococcus, Streptococcus, Treponema, Xanthomonas, Vibrio, and Yersinia.*

For therapeutic uses, the compositions or agents identified using the methods disclosed herein may be administered systemically, for example, formulated in a pharmaceutically-acceptable buffer such as physiological saline. Treatment may be accomplished directly, e.g., by treating the animal with antagonists which disrupt,

15 suppress, attenuate, or neutralize the biological events associated with a pathogenicity polypeptide (e.g., a biofilm regulator polypeptide). Preferable routes of administration include, for example, subcutaneous, intravenous, interperitoneally, intramuscular, or intradermal injections which provide continuous, sustained levels of the drug in the patient. Treatment of human patients or other animals will be carried out using a

20 therapeutically effective amount of an anti-pathogenic agent in a physiologically-acceptable carrier. Suitable carriers and their formulation are described, for example, in Remington's Pharmaceutical Sciences by E.W. Martin. The amount of the anti-pathogenic agent (e.g., an antibiotic) to be administered varies depending upon the manner of administration, the age and body weight of the patient, and with the type of

25 disease and extensiveness of the disease. Generally, amounts will be in the range of those used for other agents used in the treatment of other microbial diseases, although in certain instances lower amounts will be needed because of the increased specificity of the compound. A compound is administered at a dosage that inhibits microbial proliferation (e.g., biofilm formation). If desired, such treatment is also performed in

30 conjunction with standard antibiotic therapy.

Other Embodiments

In general, the invention includes any nucleic acid sequence which may be isolated as described herein or which is readily isolated by homology screening or PCR amplification using the nucleic acid sequences of the invention. Also included in the invention are polypeptides which are modified in ways which do not abolish their pathogenic activity (assayed, for example as described herein). Such changes may include certain mutations, deletions, insertions, or post-translational modifications, or may involve the inclusion of any of the polypeptides of the invention as one component of a larger fusion protein. Also, included in the invention are polypeptides that have lost their pathogenicity.

Thus, in other embodiments, the invention includes any protein which is substantially identical to a polypeptide of the invention. Such homologs include other substantially pure naturally-occurring polypeptides as well as allelic variants; natural mutants; induced mutants; proteins encoded by DNA that hybridizes to any one of the nucleic acid sequences of the invention under high stringency conditions or, less preferably, under low stringency conditions (e.g., washing at 2X SSC at 40°C with a probe length of at least 40 nucleotides); and proteins specifically bound by antisera of the invention.

The invention further includes analogs of any naturally-occurring polypeptide of the invention. Analogs can differ from the naturally-occurring the polypeptide of the invention by amino acid sequence differences, by post-translational modifications, or by both. Analogs of the invention will generally exhibit at least 85%, more preferably 90%, and most preferably 95% or even 99% identity with all or part of a naturally-occurring amino acid sequence of the invention. The length of sequence comparison is at least 15 amino acid residues, preferably at least 25 amino acid residues, and more preferably more than 35 amino acid residues. Again, in an exemplary approach to determining the degree of identity, a BLAST program may be used, with a probability score between e^{-3} and e^{-100} indicating a closely related sequence. Modifications include *in vivo* and *in vitro* chemical derivatization of polypeptides, e.g., acetylation, carboxylation, phosphorylation, or glycosylation; such modifications may occur during polypeptide

synthesis or processing or following treatment with isolated modifying enzymes.

Analogs can also differ from the naturally-occurring polypeptides of the invention by alterations in primary sequence. These include genetic variants, both natural and induced (for example, resulting from random mutagenesis by irradiation or exposure to ethanemethylsulfate or by site-specific mutagenesis as described in Sambrook, Fritsch and Maniatis, *Molecular Cloning: A Laboratory Manual* (2d ed.), CSH Press, 1989, or Ausubel et al., *supra*). Also included are cyclized peptides, molecules, and analogs which contain residues other than L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids.

10 In addition to full-length polypeptides, the invention also includes fragments of any one of the polypeptides of the invention. As used herein, the term "fragment," means at least 5, preferably at least 20 contiguous amino acids, preferably at least 30 contiguous amino acids, more preferably at least 50 contiguous amino acids, and most preferably at least 60 to 80 or more contiguous amino acids. Fragments of the invention
15 can be generated by methods known to those skilled in the art or may result from normal protein processing (e.g., removal of amino acids from the nascent polypeptide that are not required for biological activity or removal of amino acids by alternative mRNA splicing or alternative protein processing events).

Furthermore, the invention includes nucleotide sequences that facilitate specific
20 detection of any of the nucleic acid sequences of the invention. Thus, for example, nucleic acid sequences described herein or fragments thereof may be used as probes to hybridize to nucleotide sequences by standard hybridization techniques under conventional conditions. Sequences that hybridize to a nucleic acid sequence coding sequence or its complement are considered useful in the invention. Sequences that
25 hybridize to a coding sequence of a nucleic acid sequence of the invention or its complement and that encode a polypeptide of the invention are also considered useful in the invention. As used herein, the term "fragment," as applied to nucleic acid sequences, means at least 5 contiguous nucleotides, preferably at least 10 contiguous nucleotides, more preferably at least 20 to 30 contiguous nucleotides, and most preferably at least 40

to 80 or more contiguous nucleotides. Fragments of nucleic acid sequences can be generated by methods known to those skilled in the art.

The invention further provides a method for inducing an immunological response in an individual, particularly a human, which includes inoculating the individual with, for example, any of the polypeptides (or a fragment or analog thereof or fusion protein) of the invention to produce an antibody and/or a T cell immune response to protect the individual from infection, especially bacterial infection (e.g., a *Pseudomonas aeruginosa* infection). The invention further includes a method of inducing an immunological response in an individual which includes delivering to the individual a nucleic acid vector to direct the expression of a polypeptide described herein (or a fragment or fusion thereof) in order to induce an immunological response.

The invention also includes vaccine compositions including the polypeptides or nucleic acid sequences of the invention. For example, the polypeptides of the invention may be used as an antigen for vaccination of a host to produce specific antibodies which protect against invasion of bacteria. The invention therefore includes a vaccine formulation which includes an immunogenic recombinant polypeptide of the invention together with a suitable carrier.

The invention further provides compositions (e.g., nucleotide sequence probes), polypeptides, antibodies, and methods for the diagnosis of a pathogenic condition.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

What is claimed is:

Claims

1. An isolated polypeptide comprising an amino acid sequence having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2), wherein
5 expression of said polypeptide, in a microorganism, affects phenotype-mediated antibiotic-resistance in said microorganism.
2. The isolated polypeptide of claim 1, said polypeptide comprising the amino acid sequence of PvrR (SEQ ID NO:2).
- 10 3. The isolated polypeptide of claim 1, wherein said amino acid sequence consists essentially of the amino acid sequence of PvrR (SEQ ID NO:2) or a fragment thereof.
- 15 4. An isolated polypeptide fragment of the isolated polypeptide of claim 1.
5. The isolated polypeptide fragment of claim 4, wherein said polypeptide fragment comprises 200 contiguous amino acids of SEQ ID NO:2.
- 20 6. An isolated polynucleotide having at least 50% identity to the nucleotide sequence of *pvrR* (SEQ ID NO:1), wherein expression of said polynucleotide, in a microorganism, affects phenotype-mediated antibiotic-resistance in said microorganism.
- 25 7. The isolated polynucleotide of claim 6, said polynucleotide comprising the nucleotide sequence of *pvrR* (SEQ ID NO:1) or a complement thereof.
8. The isolated polynucleotide of claim 7, said polynucleotide consisting essentially of the nucleotide sequence of *pvrR* (SEQ ID NO:1) or a fragment thereof.

9. A vector comprising the isolated polynucleotide of any one of claims 6, 7, or 8.

10. A host cell comprising the vector of claim 9.

5

11. A screening method for identifying a compound that modulates gene expression of a regulator polynucleotide that affects phenotype-mediated antibiotic-resistance in a microorganism, said method comprising the steps of:

(a) providing a microbial cell comprising a polynucleotide having at least 50% identity to the nucleotide sequence of *pvrR* (SEQ ID NO:1), wherein expression of said polynucleotide, in said microbial cell, affects phenotype-mediated antibiotic-resistance in said microbial cell;

(b) contacting said microbial cell with a compound; and

(c) comparing the level of gene expression of said polynucleotide in the presence of said compound with the level of gene expression in the absence of said compound; wherein a measurable difference in gene expression indicates that said compound modulates gene expression of a regulator polynucleotide that affects phenotype-mediated antibiotic-resistance in a microorganism.

12. The method of claim 11, wherein said screening method identifies a compound that increases transcription of said regulator polynucleotide.

13. The method of claim 11, wherein said screening method identifies a compound that decreases transcription of said regulator polynucleotide.

25

14. The method of claim 11, wherein said screening method identifies a compound that increases translation of an mRNA transcribed from said regulator polynucleotide.

15. The method of claim 11, wherein said screening method identifies a compound that decreases translation of an mRNA transcribed from said regulator polynucleotide.

5 16. The method of claim 11, wherein the compound is a member of a chemical library.

17. The method of claim 11, wherein said microbial cell belongs to the genus *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*.

10 18. The method of claim 11, wherein said microbial cell is a phenotypic variant having increased biofilm formation.

15 19. The method of claim 18, wherein said phenotypic variant is a small colony variant.

20 20. The method of claim 19, wherein said small colony variant is a small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*.

21. The method of claim 18, wherein said small colony variant is a rough small colony variant.

22. The method of claim 21, wherein said rough small colony variant is *Pseudomonas*, *Vibrio*, or *Salmonella*.

25 23. The method of claim 11, wherein the activity of the compound is dependent upon the presence of the *pvrR* gene (SEQ ID NO:1) or a functional equivalent thereof.

24. The method of claim 11, wherein said compound targets the *pvrR* gene (SEQ ID NO:1) or a functional equivalent thereof.

25. The method of claim 11, wherein expression of said polynucleotide
5 mediates phenotypic switching of said microbial cell in the presence of a high concentration of an antibiotic.

26. The method of claim 11, wherein said polypeptide is expressed by the isolated polynucleotide of any one of claims 6, 7, or 8.
10

27. A screening method for identifying a compound that modulates an activity of a polypeptide that affects phenotype-mediated antibiotic-resistance in a microorganism, said method comprising the steps of:

- 15 (a) providing a microbial cell expressing a polypeptide having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2), wherein expression of said polypeptide, in said microbial cell, affects phenotype-mediated antibiotic-resistance in said microbial cell;
- (b) contacting said microbial cell with a compound; and
- 20 (c) comparing an activity of said polypeptide in the presence of said compound with said activity in the absence of said compound; wherein a measurable difference in the activity indicates that said compound modulates said activity of said polypeptide that affects phenotype-mediated antibiotic-resistance in a microorganism.

28. The method of claim 27, wherein said screening method identifies a
25 compound that increases the activity of said polypeptide.

29. The method of claim 27, wherein said screening method identifies a compound that decreases the activity of said polypeptide.

30. The method of claim 27, wherein the compound is a member of a chemical library.

31. The method of claim 27, wherein comparing the activity of the polypeptide involves an immunological assay.

32. The method of claim 27, wherein said microbial cell belongs to the genus *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*.

33. The method of claim 27, wherein said microbial cell is a phenotypic variant having increased biofilm formation.

34. The method of claim 33, wherein said phenotypic variant is *Pseudomonas aeruginosa* PA14 RSCV.

35. The method of claim 27, wherein said regulator polypeptide is the isolated polypeptide of claim 1.

36. The method of claim 27, wherein the activity of the polypeptide regulates phenotypic switching.

37. The method of claim 27, wherein the activity of the polypeptide regulates biofilm-mediated antibiotic-resistance.

38. The method of claim 27, wherein the activity of the polypeptide affects susceptibility of the microbial cell to antibiotic treatment.

39. The method of claim 27, wherein said polypeptide is an element of a two-component regulatory system.

40. The method of claim 27, wherein the activity of the compound is dependent upon the presence of the PvrR polypeptide (SEQ ID NO:2) or a functional equivalent thereof.

5 41. The method of claim 27, wherein said compound targets the PvrR polypeptide (SEQ ID NO:2) or a functional equivalent thereof.

42. The method of claim 27, wherein said polypeptide mediates phenotypic switching of said microbial cell in the presence of a high concentration of an antibiotic.

10

43. The method of claim 27, wherein said polypeptide is expressed by the isolated polynucleotide of any one of claims 6, 7, or 8.

44. A screening method for identifying a compound that modulates microbial biofilm formation, said method comprising the steps of:

15

(a) culturing a microbial cell comprising a polypeptide having at least 50% identity to the amino acid sequence of PvrR (SEQ ID NO:2), wherein said microbial cell, upon culturing, forms a biofilm;

(b) contacting said microbial cell with a compound; and

20

(c) comparing microbial biofilm formation in the presence of said compound with microbial biofilm formation in the absence of said compound; wherein a measurable difference in said microbial biofilm formation indicates that said compound modulates biofilm formation.

25 45. The method of claim 44, wherein said screening method identifies a compound that increases biofilm formation.

46. The method of claim 44, wherein said screening method identifies a compound that decreases biofilm formation.

30

47. The method of claim 44, wherein biofilm formation is measured by assaying microbial aggregation.

5 48. The method of claim 47, wherein microbial aggregation is assayed using a microscope.

49. The method of claim 47, wherein microbial aggregation is assayed using a salt aggregation test.

10 50. The method of claim 47, wherein microbial aggregation is assayed using an attachment assay.

51. The method of claim 44, wherein the compound is a member of a chemical library.

15

52. The method of claim 44, wherein said microbial cell belongs to the genus *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*.

53. The method of claim 44, wherein said microbial cell is a phenotypic variant having increased biofilm formation.

20

54. The method of claim 53, wherein said phenotypic variant is a small colony variant.

25 55. The method of claim 54, wherein said small colony variant is a small colony variant of *Pseudomonas*, *Vibrio*, *Salmonella*, or *Staphylococcus*.

56. The method of claim 54, wherein said small colony variant is a rough small colony variant.

30

57. The method of claim 56, wherein said rough small colony variant is *Pseudomonas*, *Vibrio*, or *Salmonella*.

58. The method of claim 44, wherein the activity of the compound is
5 dependent upon the presence of PvrR polypeptide (SEQ ID NO: 2) or a functional equivalent thereof.

59. The method of claim 44, wherein said compound targets the PvrR
polypeptide (SEQ ID NO:2) or a functional equivalent thereof.

10

60. The method of claim 44, wherein expression of said polypeptide mediates phenotypic switching of said microbial cell in the presence of a high concentration of an antibiotic.

15 61. The method of claim 44, wherein said polypeptide is an isolated polypeptide of any one of claims 1, 2, or 3.

62. A method of treating a microbial infection involving a microorganism that forms a biofilm in a mammal, said method comprising administering to said
20 mammal a therapeutically-effective amount of a compound that induces the expression of or activity of or represses the expression of or activity of the polypeptide of any one of claims 1, 2, or 3.

63. The method of claim 62, wherein said method further comprises
25 administering to said mammal a therapeutically-effective amount of an antibiotic.

64. The method of claim 62, wherein said mammal is a human.

65. The method of claim 62, wherein said human has cystic fibrosis.
30

66. The method of claim 62, wherein said human has a chronic infection.

67. The method of claim 62, wherein the said microorganism belongs to the genus *Pseudomonas*, *Vibrio*, *Salmonella* or *Staphylococcus*.

5

68. A method of cleaning or disinfecting a surface at least partially covered by a microorganism that forms a biofilm, said method comprising contacting said microorganism with a cleaning composition comprising a compound that induces the expression of or activity of or represses the expression of or activity of the polypeptide of claim 1, 2, or 3.

10

69. The method of claim 68, wherein said microorganism belongs to the genera *Pseudomonas*, *Vibrio*, *Salmonella* or *Staphylococcus*.

15

70. A screening method for identifying a compound that decreases pathogenicity of an antibiotic-resistant phenotypic variant, said method comprising the steps of:

(a) contacting an antibiotic-resistant phenotypic variant with a candidate compound; and

20

(b) measuring reversion of said antibiotic-resistant phenotypic variant to a wild-type phenotype, an increase in reversion indicating that said compound decreases pathogenicity of said antibiotic-resistant phenotypic variant.

25

71. The method of claim 70, wherein said antibiotic-resistant phenotypic variant is a bacterial variant.

72. The method of claim 71, wherein said antibiotic-resistant phenotypic bacterial variant is cultured in the absence of an antibiotic.

73. The method of claim 71, wherein said antibiotic-resistant phenotypic bacterial variant has increased biofilm formation.

74. The method of claim 71, wherein said antibiotic-resistant phenotypic bacterial variant is a rough small colony variant.

75. The method of claim 71, wherein said antibiotic-resistant phenotypic bacterial variant is a hyperpiliated variant.

76. The method of claim 71, wherein said antibiotic-resistant phenotypic bacterial variant has increased hydrophobicity.

77. The method of claim 71, wherein said antibiotic-resistant phenotypic bacterial variant has an alteration in a surface component.

78. The method of claim 71, wherein said antibiotic-resistant phenotypic bacterial variant is a pathogen.

79. The method of claim 78, wherein said pathogen is a Gram positive bacterium.

80. The method of claim 79, wherein said pathogen is *Staphylococcus*.

81. The method of claim 78, wherein said pathogen is a Gram negative bacterium.

82. The method of claim 75, wherein said pathogen is *Vibrio*, *Pseudomonas*, or *Salmonella*.

83. A screening method for identifying a compound that decreases pathogenicity of a wild-type microbe, said method comprising the steps of:

(a) culturing a wild-type microbe with a candidate compound in the presence of an antibiotic; and

5 (b) comparing the number of antibiotic-resistant phenotypic variants in the presence of said compound to the number of antibiotic-resistant phenotypic variants in the absence of said compound, a decrease in the number of said antibiotic-resistant phenotypic variants in the presence of said compound indicating that said compound decreases pathogenicity of said wild-type microbe.

10

84. A screening method for identifying a polynucleotide encoding a regulator polypeptide that modulates an antibiotic-resistant phenotype of a microorganism, said method comprising the steps of:

(a) identifying an antibiotic-resistant phenotypic variant of a microorganism comprising a first phenotype;

15 (b) mutagenizing said antibiotic-resistant phenotypic variant of said microorganism, thereby generating a mutated phenotypic variant of said microorganism; and

(c) selecting said mutated phenotypic variant of step (b) having a second phenotype, other than the first phenotype of said antibiotic-resistant phenotypic variant, wherein said second phenotype identifies a mutation in said mutated phenotypic variant of step (b); and

20 (d) using said mutation for identifying a polynucleotide encoding a regulator polypeptide that modulates an antibiotic-resistant phenotype of a microorganism.

25

85. The method of claim 84, wherein said second phenotype comprises a wild-type phenotype.

86. A screening method for identifying a polynucleotide encoding a regulator polypeptide that modulates phenotype-mediated antibiotic-resistance of a microorganism, said method comprising the steps of:

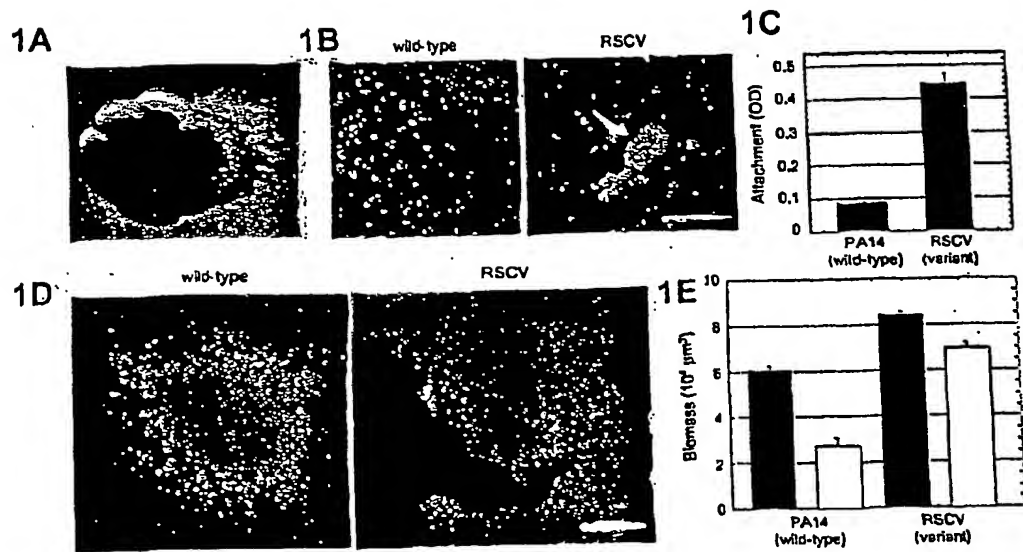
5 (a) transforming an antibiotic-resistant phenotypic variant of a microorganism with a candidate polynucleotide encoding a regulator polypeptide; and

(b) culturing said transformed antibiotic-resistant phenotypic variant of a microorganism under conditions suitable for expression of said regulator polypeptide; and

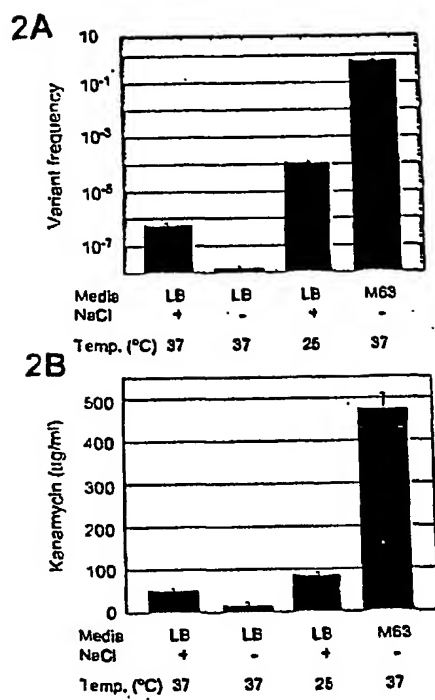
10 (c) measuring reversion of said transformed antibiotic-resistant phenotypic variant of said microorganism to a wild-type phenotype, an increase in reversion identifies said polynucleotide as encoding a regulator polypeptide that modulates phenotype-mediated antibiotic-resistance.

87. The method of claim 80, wherein said polynucleotide encodes a regulator polypeptide that modulates a phenotypic switch from antibiotic-resistant phenotype to an antibiotic-susceptible phenotype.

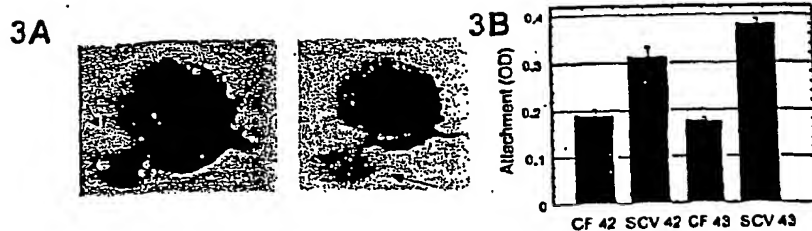
88. The method of claim 80, wherein said polynucleotide having at least 50% identity to the nucleotide sequence of *pvrR* (SEQ ID NO:1) encodes an element of a two-component regulatory system.



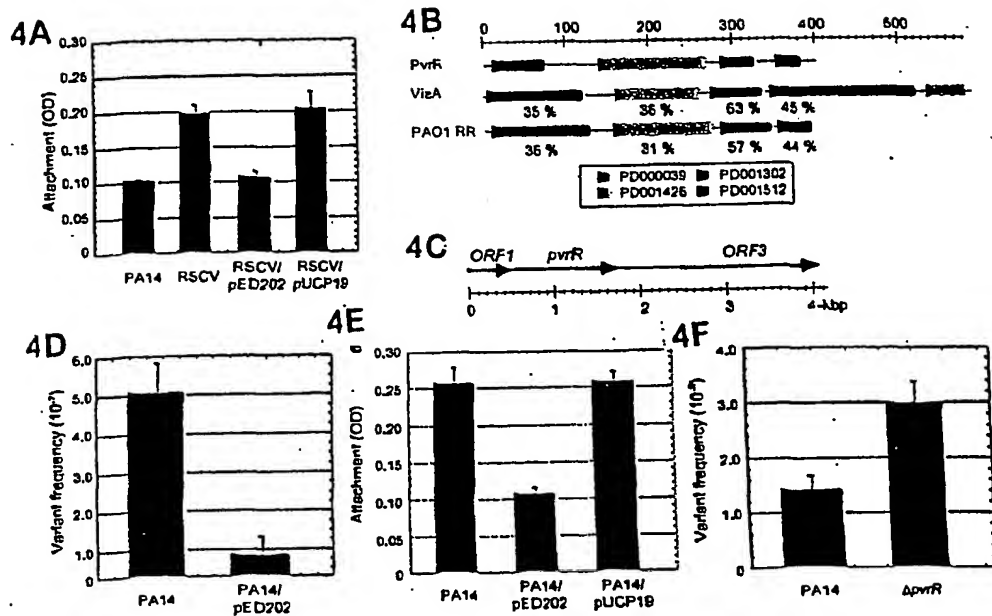
Figures 1 A-E



Figures 2 A-B



Figures 3A-B



Figures 4 A-F

Figure 5A, pvrR (SEQ ID NO:1)

```

atgagctgga aatcctatcg ggtgctggtg gtcgaagatc agccgtttca gcgcgaatac 60
ctgctcaacc tgtttcgcga gcgcggcggtg cagtacctgg taggtgccgg cgacggcgcg 120
gagggcggtgc gctgcctgaa gcaggacagg ttcgacctga tcctcagcga tctgatgatg 180
ccgggcatgg atggtatcca aatgatcctg caactgccgt atctcaagca tcgtccgaag 240
ctggcgctga tgagctcctc gtcgcagcgg atgatgctca gtgccagccg ggtcgcccag 300
agtctcggtt tgtcggtaat cgacctgttg cccaagccga ctctgcccaa ggccatcggc 360
caacttctgg aacacctgga aagatgcctc aggcagaagc tggagccgga aaccgacgag 420
actccgcatg ggcgcacggc gttgctggat gccctgcata acgagcaact ggtgacctgg 480
ttccaggcta agaaatccct ccacaccggg cgcatagtcg gcgccgaggc gttgatacgc 540
tggagccacc cgcagcatgg cctggtgctg cccagctggt tcatgagtga tgtcgacgct 600
accggtctgc acgaggcggt gctctggcgc gtgctcgaac agaccctgaa cgcccaggaa 660
tcgtggcgca gggcgggtta cgagattccg gtttcgggtg atctgccgcc gcacctgctc 720
gataaccagg aacttccgga tcgactctat gagtacgtcg gcgctcgcgg ggcttgtacc 780
agctcactat gtttcgagtt gaccgagagc agtgtcacia ctctgtcaag taactactat 840
gcagggtgct gtgcgttgcg catgaaaggg ttcggattgg cccaggacga ctttggccag 900
ggttacagct cgttctataa cctggtcacg acgcctttca cggagctgaa gatcgaccgc 960
tccctagtcc agggatgcgt agaggataac ggcctcaatg cagctgtcat cagttgtatt 1020
gagttgggtc accgcctgaa tctcgacgtg gtggccgaag gcgtggagac ctgagaggaa 1080
ctgaatcttc ttcgtcgtct tggctgcgac cgggcgcagg gtttcctgat ttctaaggca 1140
gtgtctgctc gtgagttcga gcggcagtta agggaggacg gccccagcct ccttgtttaa 1200

```

Figure 5B

ORF1-12

SEQ ID NO:3

```

atgagccatg aaatccgcac acctctgtac ggcattgctg gcacgcttga gctgcttggg 60
cgtaccgagc tgagtcggca gcaggccggt tacctaaagg caatccagca ttcctcgtcg 120
accctgctgc aactgatcag cgatgtgctt gacgtatcca agatagaggc cggccaactg 180
gacctagagt gcgtggaatt ctcccgttg gaattgaccg aagaggctcg gcagtcgttc 240
accggtgccg cgcaggccaa ggggctgcag ttgtatacct gctctctg gcagctgccg 300
ctgcgcatgc gggggggcgc ggcgtcgatc cggcagattc tcaacaacct gctgagcaac 360
gcggtgaagt tcaccgacaa tggctatgtc aacgtccacc tgaaggccag cgtggctgat 420
gccgaatgtg tgatgctgac ctggcaggtc aacgataccg gcatggggat caacgtcgag 480
gatcagccgc gtctgttcga accgttctac cagatacgcc gctccgagca tccggctcgca 540
ggcacgggccc tcggcttgct gatcagccag cgcctggcgc agctaataaa tggcagctctg 600
aaactggtca gtgagctggg gttgggcagc agctttagcc tcaggcttcc gcttgagcgg 660
atcgcgatgc aggctgagcc gcaggacct gcccgggtgc cgtccaagt gctggcgcct 720
gtccgcgacc taacggaatg cctgtgtggc tggatctccc gctggggtgg aagggccatg 780
gtcgcgacgc cgaggctcgt ggacgaggcg gacgcgacct cgtgctggt cgaagtgtta 840
ctgctggagg gggcgccgat gttcgaagca tggccaggat gccgggtgga gctttccctt 900
cagggtgata tggagccgca ggcacagggc cgcgactggc tgctcgggct caacaacctg 960
gacggcctgc atcgtgctct gggcctggcc catgggcgtc tcgctgatcc ttcgacgccg 1020
ccgatacggc tggctccgtt gcgcaatcta ggtctccgct tcctagtggg ggaggataac 1080
gcgatcaacc agttgatctt gagggaccag atggaagcgc tgggctgcag cgtggagctg 1140
ctcttcgatg gtcgcgaggg gttgctgcac tgccagacgg cctgcttcga cgtgggtgctc 1200
accgatatac acatgccgaa catgaacgga tacgagctaa ccgcggagct acggcgccaa 1260
gggttccggc agccgatcat cggcgcgacg gcgaacgcca tgcgtgagga gcgcgagcgc 1320
tgcatgtccg ccgggatgaa cgattgcctg gtcaaaccgg tggatctgaa tgcccttcag 1380
aactgcttga ttaatatctt caaggtggat cgatga 1416

```

Figure 5C ORF3 (SEQ ID NO:5)

```

atgatggatg ttatacggga gcatgaggtg tttcttgggc gcatcgctcg aaaaagcgac 60
aagaccaccc agaagtacga ctatgacgtg gtgcctttgc agcggcactt gttggcaaag 120
gaaaacggat tagcgggtcta tgagggacgg gagttttcct ttgctatgcc atttctactg 180
gctaccaagc acgcgttgag cgccgattcc tcgggagatc cgttttcgct cgggtgtattg 240
ctcgccaatt tctacggaag cttctggagt gtttccgctt atcccgcgcc acagttactg 300
atctttgatc tttccggcag caccgcctg gcagtgccgt cgattccctc cacagcgag 360
cgtgacaggt tgagcggaag ctatccgatg atagtcgagc gcattctggc gcgcttgccg 420
acccggccgg tgggggagga cgctcagcgt gtccattgga tacgcgtga tcgctatcgc 480
gactcggcgc tggagatggt gggagtcgcc cgggttgatc tgccggaaac actctggtgg 540
cacgacgagc cgaaccaatc gatcatcgct gcgagcctgc ttgatctcag gcgaatcaat 600
gacttcgaac agttggttga gcgcccggca ttcgattcgt acagcctggt atcgccggat 660
ggcgaggatg tgctcggcgc gggccctgcg accggcctga gggatggcct gaacctcacc 720
cgacaggggg tcgccgttca actgcgcagc cagcctgaga acggctggct cgcgggtctac 780
cgaaccgact acggcaattt ctttcgccac tcccggtggt tgggtggcagg tctgctgctg 840
accccgccgc tgctcctggc cggttggttc gggatgcgtt ggtacaccag cagcgtcgtc 900
aaccgggtgc atcgggcgca ccggcaactg gtggagagcg acaccttcag ccggacgctg 960
atacagaccg cgccggtggc tctggtggtg ctgaccagg atgaccagca actggtgacc 1020
tgcaaccact tggccgcccc gtggtgggcg gggcccacgg agatccttgg gctgacttcc 1080
aactggaagc ttttcgatgc gcgtgggcag gtaccaggag acatctgtat ccaggtcggt 1140
ggcgctatt tgacagaccg cttcgcggcg acccgctatg ccggcaccga ggcggtactg 1200
tgctatttca acgacatcac ggtccactgc gaggcggaga ccgcgctgtc caatgcgaag 1260
cgagcagcgg atgccgccag ccaggccaag accctgttcc tggcccgcac gagccatgaa 1320
atccgtactc ccctgtacgg tgtccttggc accctggagt tgctcgacct gaccacctg 1380
aacgagcggc aacgcgccta cctacgcacc atccagagtt cgtctgcgac gctcatgcaa 1440
ctgattagcg atgtgctgga tgtctcgaag atcgaagcgg ggcagatggc tctgaccctg 1500
gccgccttca atccgctgga cctagtgcgg gaagtgcctt gcaactttgc cgccagcgcc 1560
atggccaagg acctgcaggt agaccgcctc gatactcttg cgcttgaggc gcaggtcgcg 1620
catggcttcg aagaaagcgt tctgttcgag gttgctgggt gctcggtcgg ccatttcgaa 1680
gaggtgtcgc tcggcgttgt cgaacaacgc ctgcaacgcc tgtttcagct gcagcgccgc 1740
cttgctcgcg acctgcacga ggatgaccgg caggcgcccc gctccggcgt tcggcgacgg 1800
ctcggaagcg accctggtca ggtgcaccac attggcatcg ttctgcatcg ggactctcct 1860
gccacctcgc cggccgcgca tgggaatggc aaaatcgggc acagaggatc gattggcgctc 1920
gtccgtaacg tcaatttcca ggcgtcaaaa acaagtatct acattcatta tagagatact 1980
ttcaaatcta gatag 1995

```


Figure 5D PvrR (SEQ ID NO:2)

```

Met Ser Trp Lys Ser Tyr Arg Val Leu Val Val Glu Asp Gln Pro Phe
 1           5           10           15
Gln Arg Glu Tyr Leu Leu Asn Leu Phe Arg Glu Arg Gly Val Gln Tyr
 20           25           30
Leu Val Gly Ala Gly Asp Gly Ala Glu Ala Leu Arg Cys Leu Lys Gln
 35           40           45
Asp Arg Phe Asp Leu Ile Leu Ser Asp Leu Met Met Pro Gly Met Asp
 50           55           60
Gly Ile Gln Met Ile Leu Gln Leu Pro Tyr Leu Lys His Arg Pro Lys
 65           70           75           80
Leu Ala Leu Met Ser Ser Ser Ser Gln Arg Met Met Leu Ser Ala Ser
 85           90           95
Arg Val Ala Gln Ser Leu Gly Leu Ser Val Ile Asp Leu Leu Pro Lys
100           105           110
Pro Thr Leu Pro Lys Ala Ile Gly Gln Leu Leu Glu His Leu Glu Arg
115           120           125
Cys Leu Arg Gln Lys Leu Glu Pro Glu Thr Asp Glu Thr Pro His Gly
130           135           140
Arg Thr Ala Leu Leu Asp Ala Leu His Asn Glu Gln Leu Val Thr Trp
145           150           155           160
Phe Gln Ala Lys Lys Ser Leu His Thr Gly Arg Ile Val Gly Ala Glu
165           170           175
Ala Leu Ile Arg Trp Ser His Pro Gln His Gly Leu Leu Leu Pro Ser
180           185           190
Cys Phe Met Ser Asp Val Asp Ala Thr Gly Leu His Glu Ala Leu Leu
195           200           205
Trp Arg Val Leu Glu Gln Thr Leu Asn Ala Gln Glu Ser Trp Arg Arg
210           215           220
Ala Gly Tyr Glu Ile Pro Val Ser Val Asn Leu Pro Pro His Leu Leu
225           230           235           240
Asp Asn Gln Glu Leu Pro Asp Arg Leu Tyr Glu Tyr Val Gly Ala Arg
245           250           255
Gly Ala Cys Thr Ser Ser Leu Cys Phe Glu Leu Thr Glu Ser Ser Val
260           265           270
Thr Thr Leu Ser Ser Asn Tyr Tyr Ala Gly Ala Cys Arg Leu Arg Met
275           280           285
Lys Gly Phe Gly Leu Ala Gln Asp Asp Phe Gly Gln Gly Tyr Ser Ser
290           295           300
Phe Tyr Asn Leu Val Thr Thr Pro Phe Thr Glu Leu Lys Ile Asp Arg
305           310           315           320
Ser Leu Val Gln Gly Cys Val Glu Asp Asn Gly Leu Asn Ala Ala Val
325           330           335
Ile Ser Cys Ile Glu Leu Gly His Arg Leu Asn Leu Asp Val Val Ala
340           345           350
Glu Gly Val Glu Thr Cys Glu Glu Leu Asn Leu Leu Arg Arg Leu Gly
355           360           365
Cys Asp Arg Ala Gln Gly Phe Leu Ile Ser Lys Ala Val Ser Ala Arg
370           375           380
Glu Phe Glu Arg Gln Leu Arg Glu Asp Gly Pro Ser Leu Leu Val
385           390           395

```

Figure 5E

ORF1-12

SEQ ID NO:4

Met	Ser	His	Glu	Ile	Arg	Thr	Pro	Leu	Tyr	Gly	Met	Leu	Gly	Thr	Leu
1				5					10					15	
Glu	Leu	Leu	Gly	Arg	Thr	Glu	Leu	Ser	Arg	Gln	Gln	Ala	Gly	Tyr	Leu
			20					25				30			
Lys	Ala	Ile	Gln	His	Ser	Ser	Ser	Thr	Leu	Leu	Gln	Leu	Ile	Ser	Asp
		35					40				45				
Val	Leu	Asp	Val	Ser	Lys	Ile	Glu	Ala	Gly	Gln	Leu	Asp	Leu	Glu	Cys
	50				55					60					
Val	Glu	Phe	Ser	Pro	Leu	Glu	Leu	Thr	Glu	Glu	Val	Val	Gln	Ser	Phe
65					70				75					80	
Thr	Gly	Ala	Ala	Gln	Ala	Lys	Gly	Leu	Gln	Leu	Tyr	Thr	Cys	Leu	Ser
				85				90					95		
Ala	Glu	Leu	Pro	Leu	Arg	Met	Arg	Gly	Ala	Ala	Ala	Ser	Ile	Arg	Gln
			100					105					110		
Ile	Leu	Asn	Asn	Leu	Leu	Ser	Asn	Ala	Val	Lys	Phe	Thr	Asp	Asn	Gly
		115					120					125			
Tyr	Val	Asn	Val	His	Leu	Lys	Ala	Ser	Val	Val	Asp	Ala	Glu	Cys	Val
	130					135					140				
Met	Leu	Thr	Trp	Gln	Val	Asn	Asp	Thr	Gly	Met	Gly	Ile	Asn	Val	Glu
145					150				155					160	
Asp	Gln	Pro	Arg	Leu	Phe	Glu	Pro	Phe	Tyr	Gln	Ile	Arg	Arg	Ser	Glu
				165				170						175	
His	Pro	Val	Ala	Gly	Thr	Gly	Leu	Gly	Leu	Ser	Ile	Ser	Gln	Arg	Leu
			180					185					190		
Ala	Gln	Leu	Met	Asn	Gly	Ser	Leu	Lys	Leu	Val	Ser	Glu	Leu	Gly	Leu
		195					200					205			
Gly	Ser	Ser	Phe	Ser	Leu	Arg	Leu	Pro	Leu	Glu	Arg	Ile	Ala	Met	Gln
	210					215					220				
Ala	Glu	Pro	Gln	Asp	Leu	Ala	Gly	Cys	Ala	Val	Gln	Val	Leu	Ala	Pro
225					230				235					240	
Val	Arg	Asp	Leu	Thr	Glu	Cys	Leu	Cys	Gly	Trp	Ile	Ser	Arg	Trp	Gly
			245					250						255	
Gly	Arg	Ala	Met	Val	Ala	Thr	Pro	Arg	Ser	Leu	Asp	Glu	Ala	Asp	Ala
			260					265					270		
Thr	Ser	Leu	Leu	Val	Glu	Val	Leu	Leu	Leu	Glu	Gly	Ala	Pro	Met	Phe
	275						280					285			
Glu	Ala	Trp	Pro	Gly	Cys	Arg	Val	Glu	Leu	Ser	Pro	Gln	Gly	Asp	Met
	290					295					300				
Glu	Pro	Gln	Ala	Gln	Gly	Arg	Asp	Trp	Leu	Leu	Gly	Leu	Asn	Asn	Leu
305					310				315					320	
Asp	Gly	Leu	His	Arg	Ala	Leu	Gly	Leu	Ala	His	Gly	Arg	Leu	Ala	Asp
			325					330					335		
Pro	Ser	Thr	Pro	Pro	Ile	Arg	Leu	Ala	Pro	Leu	Arg	Asn	Leu	Gly	Leu
			340					345					350		
Arg	Val	Leu	Val	Val	Glu	Asp	Asn	Ala	Ile	Asn	Gln	Leu	Ile	Leu	Arg
		355					360					365			
Asp	Gln	Met	Glu	Ala	Leu	Gly	Cys	Ser	Val	Glu	Leu	Leu	Phe	Asp	Gly
	370					375					380				
Arg	Glu	Ala	Leu	Leu	His	Cys	Gln	Thr	Ala	Cys	Phe	Asp	Val	Val	Leu
385					390				395					400	
Thr	Asp	Ile	Asn	Met	Pro	Asn	Met	Asn	Gly	Tyr	Glu	Leu	Thr	Ala	Glu
			405					410					415		
Leu	Arg	Arg	Gln	Gly	Phe	Arg	Gln	Pro	Ile	Ile	Gly	Ala	Thr	Ala	Asn
			420					425				430			
Ala	Met	Arg	Glu	Glu	Arg	Glu	Arg	Cys	Met	Ser	Ala	Gly	Met	Asn	Asp
	435					440					445				
Cys	Leu	Val	Lys	Pro	Val	Asp	Leu	Asn	Ala	Leu	Gln	Asn	Cys	Leu	Ile

5E/1

WO 03/004689

PCT/US02/21431

450
Asn Ile Leu Lys Val Asp Arg
465 470

460

5E/2

Figure 5F, ORF3 (SEQ ID NO:6)

```

Met Met Asp Val Ile Arg Glu His Glu Val Phe Leu Gly Arg Ile Ala
 1           5           10           15
Arg Lys Ser Asp Lys Thr Thr Gln Lys Tyr Asp Tyr Asp Val Val Pro
      20           25           30
Leu Gln Arg His Leu Leu Ala Lys Glu Asn Gly Leu Ala Val Tyr Glu
      35           40           45
Gly Arg Glu Phe Ser Phe Ala Met Pro Phe Leu Leu Ala Thr Lys His
      50           55           60
Ala Leu Ser Ala Asp Ser Ser Gly Asp Pro Phe Ser Leu Gly Val Leu
65           70           75           80
Leu Ala Asn Phe Tyr Gly Ser Phe Trp Ser Val Ser Ala Tyr Pro Ala
      85           90           95
Pro Gln Leu Leu Ile Phe Asp Leu Ser Gly Ser Thr Arg Leu Ala Val
      100          105          110
Pro Ser Ile Pro Ser Thr Ala Gln Arg Asp Arg Leu Ser Gly Ser Tyr
      115          120          125
Pro Met Ile Val Glu Arg Ile Leu Ala Arg Leu Arg Thr Arg Pro Val
      130          135          140
Gly Glu Asp Ala Gln Arg Val His Trp Ile Arg Ala Asp Arg Tyr Arg
145          150          155          160
Asp Ser Ala Leu Glu Met Leu Gly Val Ala Arg Val Asp Leu Pro Glu
      165          170          175
Thr Leu Trp Trp His Asp Glu Pro Asn His Leu Ile Ile Ala Ala Ser
      180          185          190
Leu Leu Asp Leu Arg Arg Ile Asn Asp Phe Glu Gln Leu Val Glu Arg
      195          200          205
Pro Ala Phe Asp Ser Tyr Ser Leu Val Ser Pro Asp Gly Glu Val Leu
      210          215          220
Leu Gly Ala Ala Pro Ala Thr Gly Leu Arg Asp Gly Leu Asn Leu Thr
225          230          235          240
Arg Gln Gly Val Ala Val Gln Leu Arg Ser Gln Pro Glu Asn Gly Trp
      245          250          255
Leu Ala Val Tyr Arg Thr Asp Tyr Gly Asn Phe Phe Arg His Ser Arg
      260          265          270
Trp Leu Val Ala Gly Leu Leu Leu Thr Pro Ala Leu Leu Ala Gly
      275          280          285
Trp Leu Gly Met Arg Trp Tyr Thr Ser Ser Val Val Asn Pro Val His
      290          295          300
Arg Ala His Arg Gln Leu Val Glu Ser Asp Thr Phe Ser Arg Thr Leu
305          310          315          320
Ile Gln Thr Ala Pro Val Ala Leu Val Val Leu Thr Gln Asp Asp Gln
      325          330          335
Gln Leu Val Thr Cys Asn His Leu Ala Ala Gln Trp Leu Gly Gly Pro
      340          345          350
Thr Glu Ile Leu Gly Leu Thr Ser Asn Trp Lys Leu Phe Asp Ala Arg
      355          360          365
Gly Gln Val Pro Gly Asp Ile Cys Ile Gln Val Gly Gly Arg Tyr Leu
      370          375          380
Gln Thr Ala Phe Ala Ala Thr Arg Tyr Ala Gly Thr Glu Ala Val Leu
385          390          395          400
Cys Val Phe Asn Asp Ile Thr Val His Cys Glu Ala Glu Thr Ala Leu
      405          410          415
Ser Asn Ala Lys Arg Ala Ala Asp Ala Ser Gln Ala Lys Thr Leu
      420          425          430
Phe Leu Ala Arg Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Val

```

SF/1

Figure 5F Continued

435	440	445
Leu Gly Thr Leu Glu Leu Leu Asp Leu Thr Thr Leu Asn Glu Arg Gln		
450	455	460
Arg Ala Tyr Leu Arg Thr Ile Gln Ser Ser Ser Ala Thr Leu Met Gln		
465	470	475
Leu Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Met		480
485	490	495
Ala Leu Thr Leu Ala Ala Phe Asn Pro Leu Asp Leu Val Arg Glu Val		
500	505	510
Leu Gly Asn Phe Ala Ala Ser Ala Met Ala Lys Asp Leu Gln Val Asp		
515	520	525
Pro Leu Asp Thr Leu Ala Leu Glu Ala Gln Val Ala His Gly Phe Glu		
530	535	540
Glu Ser Val Leu Phe Glu Val Ala Gly Gly Ser Val Gly His Phe Glu		
545	550	555
Glu Gly Val Val Gly Val Val Glu Gln Arg Leu Gln Arg Leu Phe Gln		560
565	570	575
Leu Gln Arg Arg Leu Val Ala His Leu His Glu Asp Asp Arg Gln Ala		
580	585	590
Pro Arg Ser Gly Val Arg Arg Arg Leu Gly Ser Asp Pro Gly Gln Val		
595	600	605
His His Ile Gly Ile Val Leu His Arg Asp Ser Pro Ala Thr Leu Ala		
610	615	620
Ala Ala His Gly Met Ala Lys Ile Gly His Arg Gly Ser Ile Gly Val		
625	630	635
Val Arg Asn Val Asn Phe Gln Ala Ser Lys Thr Ser Ile Tyr Ile His		640
645	650	655
Tyr Arg Asp Thr Phe Lys Ser Arg		
660		

5F/2

Figure 5G

SEQ ID NO:7

```

cttcctcacc gcactttcat tcaccgcacc gttatcggcg tcatggacaa tgccttgcag 60
tcctgagacc tcattaggaa tcattctgata gagggaattt ttccatgccg tcagtctttc 120
ggcaggaagt tgaagaattt cttacagcct tttgatagcg gtttctccac tccgagtgtc 180
gcgctcaagc tgcctccgca gctcgggtggc gccttgaggt tgtgcgtgct atgcagcctg 240
atattcagtg tgagcaggtt tttaaacccat caggtgtccc tcagtcggca agctaggaat 300
gtggctaggt acgaagcgca gctttatttc gagcagcgcg aggcgttgct caatcacttg 360
agcggcaatg tcgtgccctt ggccgcgggt agagcgctcg tcaacgaagc gccgaacaat 420
gtgagcatcc tgccgttgag tgacggaggg cgaggtctgc tattgaccgc tcgcacgctc 480
ggtgatctcc gggaaaagcg gctggcactg aggtatctgg tcgataccga caaaggccct 540
ctggtttacc ggcttaccgc cgatggtagg cctcggcag cgatatccag cacgataacc 600
aaagagggtg accgagcctt gctggcgact ccgtcggcg cgtttcactg ggtgactgac 660
ggtggtaccc ctcaacggct gtaccttttt gaatccttag gcgatgagcc gggcgagggg 720
tggctaggcc tggagattct cggcgaagac ctcgattcga ggttgccggc gaatgatgcc 780
ggaaactacc tgcgtctgga tcagcatggg caggtcgtac tcgctacgga cgcagaggcg 840
ctggggagcg gtgcgtcgcg gacgcttttg cgtggagacg gcttcgggtt catcggtgct 900
ggcccactgc cgcagcatat ggtgcttttc cagcacgtgg ggtcttcgag ctgggatctg 960
atctatcaca tcggtatcgg tcgctgttg ctggctctgt ggctccctct gttacttgcc 1020
tctgcgttg cactcgcatg cggcacccta ctgcattggc tggtgccggag catcgagcga 1080
cgcttgatag agcccgc aaa ggcagcctt gaagcattga aggagagcga agccttttcc 1140
cgtgcagtta tccaggccgc gcccgctcg cgtgtcgctgc tgctgcgtgc cgacgccgca 1200
gtggtccttg aaaatcccca ggccgcgcaa tggctgggtg atagcgaggc gattgcccac 1260
gacgcgccga gatggatttc ccaggcgctt gcaggagggt tgaagtgttc tggagaagaa 1320
ctggaaaccg aggcagggtt acatcttcat ctcaattaca cgcaccaccg ctataacggg 1380
gaagacgtat tgttctgcgc cttcagtgaa atcagtgcac gcaagcgggt ggaggcgga 1440
ctggctcgcg caaaatccct ggcggatgct gccaatgaag ccaagacgct gtttctcgcc 1500
accctagacc atgaaatccg cacacctctg tacggcatgc ttggcacgct tgagctgctt 1560
ggcgctaccg agctgagtcg gcagcaggcc ggttacctaa aggcaatcca gcatcctctg 1620
tcgaccctgc tgcaactgat cagcgatgtg cttgacgtat ccaagataga ggccggccaa 1680
ctggacctag agtgctgga attctccccg ctggaattga ccgaagaggt cgtgcagtcg 1740
ttcaccggtg ccgcgcaggc caaggggctg cagttgtata cctgectctc tgcggagctg 1800
ccgctgcgca tgcggggggc cgcggcgctg atccggcaga ttctcaacaa cctgtgagc 1860
aacgcgggtg agttcacgga caatggctat gtcaacgata ccggcatggg gatcaacgtc 1920
gatgccgaat gtgtgatgct gacctggcag gtcaacgata ccggcatggg gatcaacgtc 1980
gaggatcagc cgcgtctgtt cgaaccgttc taccagatac gccgctccga gcatccggtc 2040
gcaggcacgg gcctcggtt gtcgatcagc cagcgctctg cgcagctaag gaatggcagt 2100
ctgaaactgg tcagttagct ggggttgggc agcagcttta gcctcaggct tccgcttgag 2160
cggatcgcg tgcaggctga gccgcaggac ctacccgggt gcgccgtcca agtgctggcg 2220
cctgtccgcg acctaacgga atgcctgtgt ggctggatct cccgctgggg tggaaaggcc 2280
atggtcgcg cgcggaggtc gctggacgag gcggacgca cctcgctgct ggtcgaagtg 2340
ttactgctgg agggggcgcc gatgttcgaa gcatggccag gatgccgggt ggagctttcc 2400
cctcagggtg atatggagcc gcaggcacag ggccgcgact ggctgctcgg gctcaacaac 2460
ctggacggcc tgcategtgc tctgggcctg gcccatgggc gtctcgctga tccttcgacg 2520
ccgccgatac ggctggctcc gttgcgcaat ctaggctctc cgttcctagt ggtggaggat 2580
aacgcgatca accagttgat cttgagggac cagatggaag cgctgggctg cagcgtggag 2640
ctgctcttcg atggtcgga ggcgttgctg cactgccaga cggcctgctt cgacgtggtg 2700
ctcaccgata tcaacatgcc gaacatgaac ggatacagac taaccgcgga gctacggcgc 2760
caagggttcc ggcagccgat catcgccgcg acggcgaacg ccatgcgtga ggagcgcgag 2820
cgctgcatgt ccgcccggat gaacgattgc ctggtcaaac cggtgatct gaatgccctt 2880
cagaactgct tgattaatat tctcaagggt gatcgagga ctggaaatcc tatcgggtgc 2940
tggtggtcga agatcagccg tttcagcgcg aatacctgct caacctgttt cgcgagcgcg 3000
gcgtgcagta cctggttagt gccggcgacg gcgcggaggc gttgcgctgc ctgaagcagg 3060
acaggttcga cctgatctc agcgatctga tgatgccggg catggatggt atccaaatga 3120
tcctgcaact gccgatctc aagcatcgtc cgaagctggc gctgatgagc tcctcgctgc 3180
agcggatgat gctcagtgcc agccgggtcg cccagagtct cggcttgctg gtaatcgacc 3240

```

5G/1

```

tggtgccccaa gccgactctg cccaaggcca tcggccaact tctggaacac ctggaaagat 3300
gcctcaggca gaagctggag ccggaaccg acgagactcc gcatgggccc acggcggtgc 3360
tggatgccct gcataacgag caactgggtga cctgggtcca ggctaagaaa tccctccaca 3420
ccggggcgcat agtcggcgcc gaggcgttga tacgctggag ccaccgcgag catggcctgt 3480
tgctgcccag ctgtttcatg agtgatgtcg acgctaccgg tctgcacgag gcgttgctct 3540
ggcgcggtgct cgaacagacc ctgaacgccc aggaatcgtg gcgcagggcg gggtacgaga 3600
ttccgggtttc ggtgaatctg ccgcccgcacc tgctcgataa ccaggaactt ccggatcgac 3660
tctatgagta cgtcggcgct cgcggggctt gtaccagctc actatgtttc gagttgaccg 3720
agagcagtggt cacaactctg tcaagtaact actatgcagg tgcctgtcgc ttgcgcatga 3780
aagggttcgg attggcccag gacgactttg gccaggggta cagctcgttc tataacctgg 3840
tcacgacgcc tttcacggag ctgaagatcg accgtccct agtccaggga tgcgtagagg 3900
ataacggcct caatgcagct gtcacagtt gtattgagtt gggtcaccgc ctgaatctcg 3960
acgtgggtggc cgaaggcgtg gagacctgcg aggaactgaa tcttcttcgt cgtcttggt 4020
gcgaccgggc gcaggggttc ctgatttcta aggcagtgct tgctcgtgag ttcgagcggc 4080
agttaaggga ggacggcccc agcctccttg ttccacgag tatccccatt atcgcgaggt 4140
cgatcgcgag caaccaccgt cagcgcaaca gtgtcgctga aggaggccgc atcccgtaaa 4200
gtctgctagt gccttgagc acgacaacaa acttttgctc aaatggacaa cctctcgca 4260
gagcctgagc atcggcctga tctgtgtggt ggtcgtgacc gtattgctgt tcagcatctg 4320
ttactgggtcg ctggggagat tggttcagga ggaggaggac aaagtctcct tccacttcac 4380
ccgttccatc gatgttatac gggagcatga ggtatttctt gggcgcatcg ctcgaaaaag 4440
cgacaagacc acccagaagt acgactatga cgtgggtgct ttgcagcggc acttggtggc 4500
aaaggaaaaac ggattagcgg tctatgaggg acgggagttt tcccttgctc ccatttct 4560
actggctacc aagcacgcgt tgagcgccga ttctcgga gatccgtttt cgtcgggtgt 4620
attgctcgcc aatttctacg gaagcttctg gaagctatcc gctatcccg cgccacagtt 4680
actgatcttt gatctttccg gcagcaccgg cctggcagtg ccgtcgatc cctccacagc 4740
gcagcgtgac aggttgagcg gaagctatcc gggatagtc gagcgcatc tggcgcgctt 4800
gcgcacccgg ccggtggggg aggacgctca gcgtgtccat tggatacgg ctgatcgcta 4860
tcgcgactcg gcgtgggag ttgtgggagt cgcccggtt gatctgcccg aaacactctg 4920
gtggcacgac gagccgaacc atctgatcat cgctgcgagc ctgcttgatc tcaggcgaat 4980
caatgacttc gaacagttgg ttgagcgcgc ggcattcgat tcgtacagcc tgggtatcgcc 5040
ggatggcgag ggtattgctcg gcgcggcccc tgcgaccggc ctgagggatg gcctgaacct 5100
caccgcagag ggggtcgccg ttcaactcg cagccagcct gagaacggct ggctcgcggt 5160
ctaccgaacc gactacggca atttctttcg ccactcccgg tggctgggtg caggtctgct 5220
gctgaccccg gcgtgctcc tggccgggtg gctcggggc cgttggtaca ccagcagct 5280
cgtcaaccgg gtgcatcggg cgcaccggca actggtggag agcgacacct tcagccggac 5340
gctgatacag accgcgcggg tggctctggt ggtgctgacc caggatgacc agcaactggt 5400
gacctgcaac cacttgcccg ccagtggtg gggcgggccc acggagatcc ttgggctgac 5460
ttccaactgg aagcttttcg atgcgcgtgg gcaggtaacca ggagacatct gtatccaggt 5520
cgggtggcgcc tatttgagca ccgccttcgc ggcgaccgcg tatgccggca ccgaggcgg 5580
actgtgcgta ttcaacgaca tcacgggtcca ctgcgagggc gagaccgcgc tgtccaatgc 5640
gaagcgagca gcggatgccg ccagccaggc caagaccctg ttctggccc gcatgagcca 5700
tgaaatccgt actccctgt acgggtgtcct tggcaccctg gagttgctcg acctgaccac 5760
cctgaacgag cggcaacgcg cctacctacg caccatccag agttcgtctg cgacgctcat 5820
gcaactgatt agcgatgtgc tggatgtctc gaagatcgaa gcggggcaga tggtctgac 5880
cctggccgcc ttcaatccgc tggacctagt gcgggaagtg cttggcaact ttgcccag 5940
cgccatggcc aaggacctgc aggttagacc gctcgatact cttgcgcttg aggcgaggt 6000
cgcgcatggc ttccaagaaa gcgttctgtt cgaggttgc acgctgcaa cgctgtttc agctgcagcg 6120
ccaagagggt gtcgtcgccg ttgtcgaaca acgctgcaa cgctgtttc agctgcagcg 6180
ccgcttgtc gcgcacctgc acgaggatga ccggcaggcg ccccgctccg gcgttcggcg 6240
acggctcgga agcgacctg gtcagggtgca ccacattggc atcgttctgc atcgggactc 6300
tctgccacc ctgcggccg cgcattggaat ggcaaaaatc gggcacagag gatcgattgg 6360
cgctgctcgt aacgtcaatt tccaggcgct atctacattc attatagaga 6410
tactttcaaa tctagaagga gggttttcct atagacatga ctctcttcac

```

56/2

Figure 6A

ORF1-1

SEQ ID NO: 8

```

atgaagtga agaatttctt acagcctttt gatagecggt tctccactcc gagtgetgcg 60
ctcaagctgc tccgcatgct cgggtggcgcc ttgatgttgt gcgtgctatg cagcctgata 120
ttcagtgtga gcatgggttt aaaccatcag gtgtccctca gtcggcaagc tatgaatgtg 180
gctatgtacg aagcgcagct ttatttcgag cagcgcgagg cgttgetcaa tcacttgagc 240
ggcaatgtcg tggccttggc cgcgggtaga gcgctcgta acgaagcgcc gaacaatgtg 300
agcatcctgc cgttgagtga cggagggcga ggtctgctat tgaccgetcg cagctcggt 360
gatctccggg aaaagcggct ggcactgatg tatctggtcg ataccgacaa aggccctctg 420
gtttaccggc ttaccgccga tggtaggccg tcggcagcga tatccagcac gataaccaa 480
gaggtgtacc gagccttgct ggcgactccg tcggcgctcg ttcactgggt gactgacgg 540
ggtacccttc aacggctgta cctttttgaa tccttaggcg atgagccggg cgaggggtgg 600
ctaggcctgg agattctcgg cgaagacctc gattcgatgt tgcgccgga tgatgccgga 660
aactacatgc tgctggatca gcatgggcag gtcgtactcg ctacggacgc agaggcgtg 720
ggagcgggtg cgtcgcgac gcttttgct ggagacggct tcggtttcat cgggtgctggc 780
ccactgccgc agcatatggt gcttttccag cacgtggggt cttcgagctg ggtctgctg 840
tatcacatcg gtatcggctg cctgttgctg gctctgtggc tccctctgtt acttgccctc 900
gcgttgccac tcgcagtcgg catcctactg cattggctgg tgcggagcat cgagcgacgc 960
ttgatagagc ccgcaaagcg acgccttgaa gcattgaagg agagcgaagc cttttcccg 1020
gcagtatatc aggcgcgccc cgtcgcgctg tgcgtgctgc gtcgtgccga cgccgcagtg 1080
gtcctggaaa atccccaggc gcgccaatgg ctgggtgata gcgagggcat tgcccacgac 1140
gcgccgagat ggattttccca ggcgttcgca ggaggtgtga agtgttctgg agaagaactg 1200
gaaaccgagg cagggctaca tcttcatttc aattacacgc ccaccgcta taacggtgaa 1260
gacgtattgt tctgcgcctt cagtgaatc agtgcacgca agcggatgga ggcggaactg 1320
gctcgcgcaa aatccctggc ggatgctgcc aatgaagcca agacgctgtt tctcgccacc 1380
atgagccatg aaatccgcac acctctgtac ggcattgctt gcaagcttga gctgcttggg 1440
cgtaccgagc tgagtggcga gcaggccgggt tacctaaagg caatccagca ttcctcgctg 1500
accctgctgc aactgatcag cgatgtgctt gacgtatcca agatagaggc cggccaactg 1560
gacctagagt gcgtggaatt ctccccgctg gaattgaccg aagaggctcg gcagtcgttc 1620
accggtgccg cgcaggccaa ggggctgcag ttgtatacct gcctctctgc ggagctgccg 1680
ctgcgcatgc gggggggcgc ggcgtcgatc cggcagattc tcaacaacct gctgagcaac 1740
gcggtgaaat tcaccgacaa tggctatgtc aacgtccacc tgaaggccag cgtggctgat 1800
gccgaatgtg tgatgctgac ctggcaggtc aacgataccg gcatggggat caacgtcgag 1860
gatcagccgc gtctgttcga accgttctac cagatacgcc gctccgagca tccggtcgca 1920
ggcacggggc tcggcttgct gatcagccag cgctggcgcc agctaataaa tggcagctg 1980
aaactggtca gtgagctggg gttgggcagc agcttttagc tcaggcttcc gcttgagcgg 2040
atcgcgatgc aggtgagcc gcaggacctc gccgggtgcg ccgtccaagt gctggcgctc 2100
gtccgcgacc taacggaatg cctgtgtggc tggatctccc gctgggggtg aggggccatg 2160
gtcgcgacgc cgaggtcgct ggacgaggcg gacgcgacct cgctgctggt cgaagtgtta 2220
ctgctggagg gggcgccgat gttcgaagca tggccaggat gccgggtgga gctttccctc 2280
cagggtgata tggagccgca ggcacagggc cgcgactggc tgctcgggct caacaacctg 2340
gacggcctgc atcgtgctct gggcctggcc catgggcgtc tcgctgatcc ttcgacgcgc 2400
ccgatacggc tggctccgtt gcgcaatcta ggtctcccg tcctagtggg ggaggataac 2460
gcgatcaacc agttgatctt gagggaccag atggaagcgc tgggctgcag cgtggagctg 2520
ctcttcgatg gtcgcgagge gttgctgcac tgccagacgg cctgcttcga cgtgggtgct 2580
accgatatca acatgccgaa catgaacgga tacgagctaa ccgcggagct acggcgccaa 2640
gggttccggc agccgatcat cggcgcgacg gcgaacgcca tgcgtgagga gcgcgagcgc 2700
tgcattgctc cggggatgaa cgattgcctg gtcaaaccgg tggatctgaa tgccttctag 2760
aactgcttga ttaatatctc caagtggtat cgatga 2796

```


Figure 6B

ORF1-2

SEQ ID No:9

```

atgctcgggtg ggccttgat gttgtgcgtg ctatgcagcc tgatattcag tgtgagcatg 60
gttttaaaacc atcaggtgtc cctcagtcgg caagctatga atgtggctat gtacgaagcg 120
cagctttatt tcgagcagcg cgaggcgttg ctcaatcact tgagcggcaa tgtcgtgcc 180
ttggccgcgg gtagagcgct cgtcaacgaa gcgccgaaca atgtgagcat cctgccgttg 240
agtgcggag ggcgaggtct gctattgacc gctcgcacgc tcggtgatct ccgggaaaag 300
cggctggcac tgatgtatct ggtcgatacc gacaaaggcc ctctggttta ccggcttacc 360
gccgatggta ggcctcggc agcgatatcc agcacgataa ccaaagaggt gtaccgagcc 420
ttgctggcga ctccgtcggc gcctgttcac tgggtgactg acggtggtag cctcaacgg 480
ctgtaccttt ttgaatcctt aggcgatgag ccgggcgagg ggtggctagg cctggagatt 540
ctcggcgaag acctcgattc gatgttgccg cggaatgat ccggaacta catgctgtg 600
gatcagcatg ggcaggtcgt actcgtacg gacgcagagg cgctggggag cgtgctgtg 660
cgagcgttt tgcgtggaga cggcttcggt ttcacgcgtg ctggccact gccgcagcat 720
atggtgcttt tccagcacgt ggggtcttcg agctgggac tgatctatca catcggtatc 780
ggtcgcctgt tgcgtgctct gtggctccct ctgttacttg cctctgcgtt ggcactcgca 840
gtcggcatcc tactgcattg gctgggtcgg agcatcgagc gacgcttgat agagcccga 900
aagcgacgcc ttgaagcatt gaaggagagc gaagcctttt cccgtgcagt tatccaggcc 960
gcgccgctcg cgtgtgctg gctgcgtcgt gccgacgcc cagtggctct ggaaaatccc 1020
caggcgcgcc aatggctggg tgatagcgag gcgattgccc acgacgcgcc gagatggatt 1080
tcccaggcgt tcgcaggagg tgtgaagtgt tctggagaag aactggaaac cgaggcagg 1140
ctacatcttc atctcaatta cacgcccacc cgctataacg gtgaagacgt attgttctgc 1200
gccttcagtg aaatcagtc acgcaagcgg atggaggcgg aactggctcg cgaaaaatcc 1260
ctggcggatg ctgccaatga agccaagac ctgtttctcg ccaccatgag ccatgaaatc 1320
cgacacacct tgcacggcat gcttgccacg cttgagctgc ttgggcgtac cgagctgagt 1380
cggcagcagg ccggttacct aaaggcaatc cagcattcct cgtcgaccct gctgcaactg 1440
atcagcgatg tgcttgacgt atccaagata gaggccggcc aactggacct agagtgcgtg 1500
gaattctccc cgctggaatt gaccgaagag gtcgtgcagt cgttcaccgg tgccgcgcag 1560
gccaaagggc tgcagttgta taccctgcctc aaactgctga gcaacgcggt gaagtccacc 1620
gccgcggcgt cgatccggca gattctcaac aacctgctga tccgatgccga atgtgtgatg 1680
gacaatggct atgtcaacgt ccacctgaag gccagcgtgg tcgatgccga atgtgtgatg 1740
ctgacctggc aggtcaacga taccggcatg gggatcaacg tcgaggatca gccgcgtctg 1800
ttcgaaccgt tctaccagat acgcccgtcc gagcatccgg tgcagggcac gggcctcggc 1860
ttgtcgatca gccagcgctt tagcctcagg ctcccgcttg agcggatcgc gatgcaggct 1920
ctgggggttg gcagcagctt tagcctcagg ctcccgcttg agcggatcgc gatgcaggct 1980
gagccgcagg acctagccgg gtgcgccgct caagtgtgg cgccgtgccc cgacctaacg 2040
gaatgcctgt gtggctggat ctcccgttg ggtggaagg ccattggctgc gacgccagg 2100
tcgctggacg aggcggacgc gacctcgtg ctggtcgaag tgttactgct ggagggggcg 2160
ccgatgttcg aagcatggcc aggatgccgg gtggagcttt cccctcaggg tgatatggag 2220
ccgaggcac agggccgcga ctggctgctc gggctcaaca acctggacgg cctgcacgt 2280
gctctgggcc tggcccatgg gcgtctcgct gatccttcga cgccgccgat acggctggct 2340
ccgttgcgca atctaggtct ccgcgtccta gtggtggagg ataacgcgat caaccagttg 2400
atcttgaggg accagatgga agcgtcgggc tgcagcgtgg agctgctctt cgatggctgc 2460
gaggcgttgc tgcaactgcca gacggcctgc ttcgacgtgg tgctcaccga tatcaacatg 2520
ccgaacatga acggatacga gctaaccgcg gagctacggc gccaaaggtt ccggcagccg 2580
atcatcgcg cgacggcgaa cgccatgcgt gaggagcgc agcgtgcac gtccgcggg 2640
atgaacgatt gcctgggtcaa accggtggat ctgaatgcc ttcagaactg cttgattaat 2700
attctcaagg tggatcgatg a 2721

```

Figure 6C

ORF1-3

SEQ ID NO:10

```

atgttggtgcg tgctatgcag cctgatattc agtgtgagca tggtttttaa ccatcaggtg 60
tccctcagtc ggcaagctat gaatgtggct atgtacgaag cgcagcttta ttccgagcag 120
cgcgaggcgt tgctcaatca cttgagcggc aatgtcgtgc ccttgccgcg gggtagagcg 180
ctcgtcaacg aagcgccgaa caatgtgagc atcctgccgt tgagtacggg agggcgaggt 240
ctgctattga ccgctcgcac gctcggtgat ctccgggaaa agcggctggc actgatgtat 300
ctggctgata ccgacaaagg ccctctggtt taccggctta ccgccgatgg taggccctcg 360
gcagcgatat ccagcacgat aaccaaagag gtgtaccgag ccttgctggc gactccgtcg 420
gcgcctgttc actgggtgac tgacgggtgg accctcaac ggctgtacct ttttgaatcc 480
ttaggcgatg agccgggcca ggggtggcta ggcctggaga ttctcgccga agacctcgat 540
tcgatgttgc gccggaatga tgccggaaac tacatgctgc tggatcagca tgggcaggtc 600
gtactcgcta cggacgcaga ggcgctgggg agcgggtgcgt cgcggacgct tttgcgtgga 660
gacggcttcg gtttcatcgg tgctggccca ctgccgcagc atatggtgct tttccagcac 720
gtggggtctt cgagctggga tctgatctat cacatcggtt tcggctgcct gttgctggct 780
ctgtggctcc ctctgttact tgctctgctg ttggcactcg cagtcggcat cctactgcat 840
tggctggtgc ggagcatcga gcgacgcttg atagagcccg caaagcgacg ccttgaagca 900
ttgaaggaga gcgaagcctt ttcccgtgca gttatccagg ccgcgcccggt cgcgctgtgc 960
gtgctgcgtc gtgccgacgc cgcagtggtc ctggaaaatc ccagggcgcg ccaatggctg 1020
gggtgatagc aggcgattgc ccacgacgcg ccgagatgga ttcccaggc gttcgagga 1080
gggtgtgaagt gttctggaga agaactggaa accgaggcag ggctacatct tcatctcaat 1140
tacacgcca cccgctataa cgggtgaagac gtattgttct gcgccttcag tgaaatcagt 1200
gcacgcaagc ggatggaggg ggaactggct cgcgcaaaat ccctggcgga tgctgccaat 1260
gaagccaaga cgctgtttct cgccaccatg agccatgaaa tccgcacacc tctgtacggc 1320
atgcttgcca cgcttgagct gcttgggctg accgagctga gtcggcagca ggccggttac 1380
ctaaaggcaa tccagcattc ctgctcgacc ctgctgcaac tgatcagcga tgtgcttgac 1440
gtatccaaga tagaggccgg ccaactggac ctagagtgcg tggaaattct cccgctggaa 1500
ttgaccgaag aggtcgtgca gtcgttcacc ggtgccgcgc aggccaaagg gctgcagttg 1560
tatacctgcc tctctgcgga gctgccgctg cgcattgcgg gggccgcggc gtcgatccgg 1620
cagattctca acaacctgct gagcaacgcg gtgaagtcca ccgacaatgg ctatgtcaac 1680
gtccacctga aggccagcgt ggtcgatgcc gaatgtgtga tgctgacctg gcaggtcaac 1740
gataccggca tggggatcaa cgtcgaggat cagccgcgtc tggtcgaacc gttctaccag 1800
atacgcgcgt ccgagcatcc ggtcgaggc acgggcctcg gcttgctgat cagccagcgc 1860
ctggcgagc taatgaatgg cagtctgaaa ctggctcagt agctggggtt gggcagcagc 1920
tttagcctca ggcttccgct tgagcggatc gcgatgcagg ctgagccgca ggacctagcc 1980
gggtgcgcgg tccaagtgtt ggcgcctgtc cgcgacctaa cggaatgcct gtgtggctgg 2040
atctcccgtt ggggtggaag ggccatggtc cgcacgccga ggtcgctgga cgaggcggac 2100
gcgacctcgc tgctgggtcga agtgttactg ctggaggggg cgccgatgtt cgaagcatgg 2160
ccaggatgcc ggggtggagct ttcccctcag ggtgatatgg agccgcaggc acagggccgc 2220
gactggctgc tcgggtccta caacctggac ggcctgcata gtgctctggg cctggcccat 2280
gggctgtctc ctgatccttc gacgccgcgg atacggctgg ctccgttgcg caatctaggt 2340
ctccgcgtcc tagtggtgga ggataacgcg atcaaccagt tgatcttgag ggaccagatg 2400
gaagcgctgg gctgcagcgt ggagctgtct ttcatgggtc gcgaggcggt gctgactgc 2460
cagacggcct gcttcgacgt ggtgctcacc gatatcaaca tgccgaacat gaacggatac 2520
gagctaaccg cggagctacg gcgccaaagg ttccggcagc cgatcatcgg cgcgacggcg 2580
aacgccatgg gtgaggagcg cgagcgctgc atgtccgcgg ggatgaacga ttgcctggtc 2640
aaaccgggtg atctgaatgc ccttcagaac tgcttgatta atattctcaa ggtggatcga 2700
tga

```

Figure 6D

ORF1-4

SEQ ID NO:11

```

atgggttttaa accatcaggt gtccctcagt cggcaagcta tgaatgtggc tatgtacgaa 60
gcgcagcttt atttcgagca gcgcgagggc ttgctcaatc acttgagcgg caatgtcgtg 120
cccttgcccg cgggtagagc gctcgtcaac gaagcgccga acaatgtgag catcctgccg 180
ttgagtgcag gagggcgagg tctgctattg accgctcgca cgctcgggtga tctccgggaa 240
aagcggtggt cactgatgta tctggtcgat accgacaaag gccctctggt ttaccgggtt 300
accgcccgat gtaggccctc ggcagcgata tccagcacga taaccaaaga ggtgtaccga 360
gccttgctgg cgactccgtc ggcgcctgtt cactgggtga ctgacgggtg taccctcaa 420
cggctgtacc tttttgaatc cttaggcgat gagccggggc aggggtgggt aggcctggag 480
attctcggcg aagacctcga ttcgatgttg cgccggaaat atgccggaaa ctacatgctg 540
ctggatcagc atggggcagg ctactcgtc agcgacgcag aggcgctggg gagcgggtgcg 600
tcgcccagcg ttttgctgg agacggcttc ggtttcatcg gtgctggccc actgccgcag 660
catatgggtg ttttccagca cgtgggggtct tcgagctggg atctgatcta tcacatcggt 720
atcggtcgcc tgttgctggc tctgtggctc cctctgttac ttgctctgct gttggcactc 780
gcagtcggga tctactgca ttggctgggt cggagcatcg agcgacgctt gatagagccc 840
gcaaagcgac gccttgaagc attgaaggag agcgaaagcct tttcccggtc agttatccag 900
gccgcgcccc tcgcgctgtg cgtgctgcgt cgtgcgcagc ccgcagtggt cctggaaaat 960
ccccaggcgc gccaatggct ggggtgatagc gaggcgattg cccacgacgc gccgagatgg 1020
atctccaggc cgttcgcagg aggtgtgaag tgttctggag aagaactgga aaccgaggca 1080
gggctacatc ttcattctcaa ttacacgccc acccgctata acggtgaaga cgtattgttc 1140
tgcccttca gtgaaatcag tgcacgcaag cggatggagg cggaactggc tcgcgcaaaa 1200
tccctggcgg atgctgccc aaagccaaag acgctgtttc tcgccaccat gagccatgaa 1260
atccgcacac ctctgtacgg catgcttggc acgcttgagc tgcttgggag taccgagctg 1320
agtcggcagc aggcgggtta cctaaaggca atccagcatt cctcgtcgac cctgctgcaa 1380
ctgatcagcg atgtgcttga cgtatccaa atagaggccg gccaaactgga cctagagtgc 1440
gtggaattct ccccgctgga attgaccgaa gaggtcgtgc agtcgttcac cgggtgccgcg 1500
caggccaagg ggctgcagtt gtataacctgc ctctctgcgg agctgccgct gcgcattgcg 1560
ggggccgcgg cgtcgatccg gcagattctc aacaacctgc tgagcaacgc ggtgaagttc 1620
accgacaatg gctatgtcaa cgtccacctg aaggccagcg tggctgatgc cgaatgtgtg 1680
atgctgacct ggcaggtcaa cgataccggc atggggatca acgctcagga tcagccgcgt 1740
ctgttcgaac cgttctacca gataccgcgc tccgagcatc cggctcgcagg cacgggcctc 1800
ggcttgctga tcagccagcg cctggcgagc ctaatgaatg gcagctctgaa actggtcagt 1860
gagctggggg tgggcagcag ctttagcctc aggcctccgc ttgagcggat cgcgatgcag 1920
gctgagccgc aggcctagc cgggtgcgcc gtccaagtgc tggcgccctgt ccgcgacctc 1980
acggaatgcc tgtgtggctg gatctcccgc tgggggtggaa gggccatggt cgcgacgccg 2040
aggctcgctg acgagggcga cgcgacctcg ctgctgggtc aagtgttact gctggagggg 2100
gcgcccagtg tcgaagcatg gccaggatgc cgggtggagc tttccctca ggggtgatag 2160
gagccgcagg cacagggccg cgactggctg ctccgggtca acaacctgga cggcctgcat 2220
cgtgctctgg gcctggccca tgggcgtctc gctgatcctt cgacgccgccc gatacggctg 2280
gctccgttgc gcaatctagg tctccgcgtc ctagtgggtg aggataacgc gatcaaccag 2340
ttgatcttga gggaccagat ggaagcgtcg ggctgcagcg tggagctgct cttcgatggg 2400
cgcgagggct tgctgcactg ccagacggcc tgcttcgacg tgggtgctcac cgatatcaac 2460
atgccgaaca tgaacggata cgagctaacc gcggagctac ggcgccaagg gttccggcag 2520
ccgatcatcg gcgcgacggc gaacgccatg cgtgaggagc gcgagcgctg catgtccgcc 2580
gggatgaacg attgcctggg caaacgggtg gatctgaatg cccttcagaa ctgcttgatt 2640
aatattctca aggtggatcg atga 2664

```

Figure 6E

ORF1-5

SEQ ID NO:12

```

atgaatgtgg ctatgtacga agcgcagcctt tatttcgagc agcgcgagggc gttgctcaat 60
cacttgagcg gcaatgtcgt gcccttggcc gcgggtagag cgctcgtcaa cgaagcgccg 120
aacaatgtga gcatcctgcc gttgagtgac ggagggcgag gtctgctatt gaccgctcgc 180
acgctcggtg atctccggga aaagcggctg gcactgatgt atctggtcga taccgacaaa 240
ggccctcttg tttaccggct taccgccgat ggtaggccct cggcagcgat atccagcacg 300
ataaccaaag aggtgtaccg agccttgctg gcgactccgt cggcgcctgt tcactgggtg 360
actgacggtg gtaccctca acggctgtac ctttttgaat ccttaggcga tgagccgggc 420
gaggggtggc taggcctgga gattctcggc gaagacctcg attcgatgtt gcgccggaat 480
gatgccgga actacatgct gctggatcag catgggcagg tcgtactcgc tacggacgca 540
gaggcgctgg ggagcggtgc gtcgcgagc cttttgcgtg gagacggctt cggtttcatc 600
gggtgctggcc cactgccgca gcatatggtg cttttccagc acgtgggggtc ttcgagctgg 660
gatctgatct atcacatcgg tatcggtcgc ctgttgctgg ctctgtggct ccctctgtta 720
cttgccctcg cgttggcact cgcagtcggc atcctactgc attggctggg gcggagcatc 780
gagcgacgct tgatagagcc cgcaaagcga cgcttgaag cattgaagga gagcgaagcc 840
ttttcccgctg cagttatcca ggccgcgccc gtcgcgctgt gcgtgctgcg tcgtgccgac 900
gccgcagtgg tccctgaaaa tccccaggcg cgcaaatggc tgggtgatag cgaggcgatt 960
gccacgacg cgccgagatg gatttcccag gcgttcgcag gaggtgtgaa gtgttctgga 1020
gaagaactgg aaaccgagggc agggctacat cttcatctca attacacgcc caccgctat 1080
aacggtgaag acgtattggt ctgcgccttc agtgaatca gtgcacgcaa gcggatggag 1140
gcggaactgg ctgcgcgcaa atccctggcg gatgtgccaa atgaagccaa gacgctgttt 1200
ctgcgccacca tgagccatga aatccgcaca cctctgtacg gcatgcttgg cacgcttgag 1260
ctgcttgggc gtaccgagct gagtcggcag caggccgggt acgtatccaa gatagaggcc 1380
tccctcgtcga cctgtctgca actgatcagc gatgtgcttg acgtatccaa gatagaggcc 1380
ggccaactgg acctagagtg cgtggaattc tccccgctgg aattgaccga agaggctcgtg 1440
cagtcgttca ccggtgccgc gcaggccaaag gggctgcagt tgtatacctg cctctctgcg 1500
gagctgccgc tgcgcgatgcg gggggcccgcg gcgtcgatcc ggcagattct caacaacctg 1560
ctgagcaacg cgggtgaagtt caccgacaat ggctatgtca acgtccacct gaaggccagc 1620
gtggtcgatg ccgaatgtgt gatgctgacc tggcaggtca acgataccgg catggggatc 1680
aacgtcgagg atcagccgcg tctgttcgaa ccgttctacc agatacgccg ctccgagcat 1740
ccggtcgcag gcacgggcct cggcttgtcg atcagccagc gcctggcgca gctaataaat 1800
ggcagtcctga aactggtcag tgagctgggg ttgggcagca gctttagcct caggcttccg 1860
cttgagcgga tcgcatgca ggctgagccg caggacctag ccgggtgcgc cgtccaagtg 1920
ctggcgccctg tccgcgacct aacggaatgc ctgtgtggct ggatctcccg ctgggggtgga 1980
agggccatgg tcgcgacgcc gaggtcgctg gacgaggcgg acgcgacctc gctgctggtc 2040
gaagtgttac tgctggaggg ggcccgcatg ttcgaagcat ggccaggatg ccgggtggag 2100
ctttccctc agggtgatat ggagccgcag gcacagggcc gcgactggct gctcgggctc 2160
aacaacctgg acggcctgca tcgtgctctg ggccctggccc atgggcgtct cgctgaccc 2220
tcgacgccgc cgatacggct ggctccgttg cgcaatctag gtctccgcgt cctagtgggtg 2280
gaggataacg cgatcaacca gttgatcttg agggaccaga tggaaagcgt gggctgcagc 2340
gtggagctgc tcttcgatgg tcgagggcg ttgctgcact gccagacggc ctgcttcgac 2400
gtggtgctca ccgatataaa catgccgaac atgaacggat acgagctaac cgcggagcta 2460
cggcgccaag ggttccggca gccgatcatc ggcgcgacgg cgaacgccat gcgtgaggag 2520
cgcgagcgct gcatgtccgc cgggatgaac gattgcctgg tcaaaccggg ggatctgaat 2580
gcccttcaga actgcttgat taatattctc aaggtggatc gatga 2625

```

Figure 6F

ORF1-6

SEQ ID NO:13

```

atgtacgaag cgcagcttta ttctgagcag cgcgaggcgt tgctcaatca cttgagcggc 60
aatgtcgtgc ccttgccgcg gggtagagcg ctgctcaatg aagcgcgcga caatgtgagc 120
atcctgccgt tgagtgcagg agggcgagggt ctgctattga ccgctcgcac gctcgggtgat 180
ctccgggaaa agcggctggc actgatgtat ctggctgata ccgacaaagg ccctctgggt 240
taccggctta ccgccgatgg taggcctcgc gcagcgatat ccagcacgat aaccaagag 300
gtgtaccgag ccttgctggc gactccgtcg gcgcctgttc actgggtgac tgacgggtgg 360
acctctcaac ggtgtgacct ttttgaatcc ttaggcgatg agccgggcga ggggtggcta 420
ggcctggaga ttctcggcga agacctcgat tcgatgttgc gccggaatga tgccggaaac 480
tacatgctgc tggatcagca tgggcaggct gtagctcgta cggacgcaga ggcgctgggg 540
agcgggtgcg cgcggacgct ttgctgga gacggcttcg gtttcatcgg tgctggccca 600
ctgccgcagc atatggtgct tttccagcac gtggggtctt cgagctggga tctgatctat 660
cacatcggta tcggtcgcct gttgctggct ctgtggctcc ctctgttact tgccctctgcg 720
ttggcactcg cagtccgcat cctactgcat tggctggtgc ggagcatcga gcgacgcttg 780
atagagcccc caaagcgacg ccttgaagca ttgaaggaga gcgaagcctt ttcccgtagc 840
gttatccagg ccgcgcccggt cgcgctgtgc gtgctgcgtc gtgccgacgc cgcagtgggtc 900
ctgggaaaatc cccaggcgcg ccaatggctg ggtgatagcg aggcgattgc ccacgacgcg 960
ccgagatgga tttcccaggc gttcgcagga ggtgtgaagt gttctggaga agaactggaa 1020
accgagggcag ggctacatct tcacttcaat tacacgcccc cccgctataa cggatgaagac 1080
gtattgttct gcgccttcag tgaaatcagt gtagcgaagc ggatggaggc ggaactggct 1140
cgcgcaaaat ccctggcgga tgctgccaat gaagccaaga cgctgtttct cgccaccatg 1200
agccatgaaa tccgcacacc tctgtacggc atgcttggca cgcttgagct gcttgggcgt 1260
accgagctga gtcggcagca ggcgggttac cttaaaggcaa tccagcattc ctgctcgacc 1320
ctgctgcaac tgatcagcga tgtgcttgac gtatccaaga tagaggccgg ccaactggac 1380
ctagagtgcg tggaaattctc cccgctggaa ttgaccgaag aggtcgtgca gtcgttcacc 1440
ggtgccgcgc aggccaaagg gctgcagttg tatacctgcc tctctgcgga gctgcccgtg 1500
cgcagtcggg gggccgcggc gtcgatccgg cagattctca acaacctgct gagcaacgcg 1560
gtgaagtcca ccgacaatgg ctatgtcaac gtccacctga aggccagcgt ggtcgatgcc 1620
gaatgtgtga tgctgacctg gcaggtcaac gataccggca tggggatcaa cgtcgaggat 1680
cagccgcgtc tgttcgaacc gttctaccag atacgcgcgt ccgagcatcc ggtcgcaggc 1740
acgggcctcg gcttgctgat cagccagcgc ctggcgcagc taatgaatgg cagtctgaaa 1800
ctggtcagtg agctggggtt gggcagcagc tttagcctca ggcttccgct tgagcggatc 1860
gcgatgcagg ctgagccgca ggacctagcc ggtgcgcccg tccaagtgtc ggcgcctgtc 1920
cgcgacctaa cggaatgcct gtgtggctgg atctcccgtc ggggtggaag ggccatggct 1980
gcgacgccga ggtcgctgga cgaggcggac gcgacctcgc tgctggctga agtggtactg 2040
ctggaggggg cgccgatggt cgaagcatgg ccaggatgcc ggggtggagct ttcccctcag 2100
gggtgatagg agccgcaggc acagggccgc gactggctgc tcgggctcaa caacctggac 2160
ggcctgcacg gtgctctggg cctggcccat gggcgtctcg ctgaccttc gacgcgcgcg 2220
atacggctgg ctccgttgcg caatctaggt ctccgcgtcc tagtggtgga ggataacgcg 2280
atcaaccagt tgatcttgag ggaccagatg gaagcgtggt gctgcagcgt ggagctgctc 2340
ttcgatggtc gcgaggcgtt gctgcactgc cagacggcct gcttcgacgt ggtgctcacc 2400
gatatcaaca tgccgaacat gaacggatac gagctaaccg cggagctacg gcgccaaggg 2460
ttccggcagc cgatcatcgg cgcgacggcg aacgccatgc gtgaggagcg cgagcgtgc 2520
atgtccgcgg ggatgaacga ttgcctggtc aaaccggtgg atctgaatgc ccttcagaac 2580
tgcttgatta atattctcaa ggtggatcga tga 2613

```

Figure 6G

ORF1-7

SEQ ID NO:14

```

atgtatctgg tcgataccga caaaggccct ctggtttacc ggcttaccgc cgatggtagg 60
ccctcggcag cgatatccag cacgataacc aaagaggtgt accgagcctt gctggcgact 120
ccgtcggcgc ctgttcactg ggtgactgac ggtggtaccc ctcaacgget gtaccttttt 180
gaatcccttag gcgatgagcc gggcgagggg tggctaggcc tggagattct cggcgaagac 240
ctcgattcga tgttgccgcc gaatgatgcc ggaaactaca tgctgctgga tcagcatggg 300
caggtcgtac tcgctacgga cgcagaggcg ctggggagcg gtgcgtcgcg gacgcttttg 360
cgtggagacg gcttcgggtt catcgggtgt ggcccactgc cgcagcatat ggtgcttttc 420
cagcacgtgg ggtcttcgag ctgggatctg atctatcaca tcggtatcgg tcgcctgttg 480
ctggctctgt ggctccctct gttacttgcc tctgcgttgg cactcgcagt cggcatccta 540
ctgcattggc tgggtgcggag catcgagcga cgcttgatag agcccgcaaa gcgacgcctt 600
gaagcattga aggagagcga agccttttcc cgtgcagtta tccaggccgc gcccgtcgcg 660
ctgtgcgtgc tgcgtcgtgc cgacgcccga gtggtcctgg aaaatcccca ggccgcgcaa 720
tggtcgggtg atagcgaggg gattgcccac gacgcgccga gatggatttc ccaggcgctt 780
gcaggaggtg tgaagtgttc tggagaagaa ctggaaaccg aggcagggct acatcttcat 840
ctcaattaca cgcccacccg ctataacggt gaagacgtat tgttctgcgc cttcagtga 900
atcagtgcac gcaagcggat ggaggcggaa ctggctcgcg caaaatccct ggccggtatg 960
gccaatgaag ccaagagcgt gtttctcgcc accatgagcc atgaaatccg cacacctctg 1020
tacggcatgc ttggcacgct tgagctgctt gggcgtaacc agctgagtcg gcagcaggcc 1080
ggttacctaaggcaatcca gcatctctcg tcgacctgc tgcaactgat cagcagtggt 1140
cttgacgtat ccaagataga ggccggccaa ctggacctag agtgctgga attctccccg 1200
ctggaattga ccgaagaggt cgtgcagtcg ttaccgggtg ccgcgcaggg caaggggctg 1260
cagttgtata cctgcctctc tgcggagctg ccgctgcgca tgcggggggc cgcggcgctg 1320
atccggcaga ttctcaacaa cctgctgagc aacgcggtga agttcaccga caatggctat 1380
gtcaacgtcc acctgaaggc cagcgtgggt gatgccgaat gtgtgatgct gacctggcag 1440
gtcaacgata ccggcatggg gatcaacgtc gaggatcagc cgcgtctggt cgaaccgttc 1500
taccagatac gccgctccga gcatccggtc gcaggcacgg gcctcggctt gtcgatcagc 1560
cagcgcctgg cgcagctaat gaatggcagt ctgaaactgg tcagttagct ggggttgggc 1620
agcagcttta gcctcaggct tccgcttgag cggatcgca tgcaggctga gccgcaggac 1680
ctagccgggt gcgccgtcca agtgctggcg cctgtccgag acctaaccga atgcctgtgt 1740
ggctggatct cccgctgggg tggaaaggcc atggtcgca cgcagaggtc gctggacgag 1800
gcggacgcga cctcgctgct ggtcgaagtg ttactgctgg agggggcgcc gatgttcgaa 1860
gcatggccag gatgccgggt ggagctttcc cctcagggtg atatggagcc gcaggcacag 1920
ggccgcgact ggctgctcgg gctcaacaac ctggacggcc tgcacgtgc tctgggctg 1980
gcccattggc gtctcgctga tccctcgacg ccgcccatac ggctggctcc gttgcgcaat 2040
ctaggctctc gcgtcctagt ggtggaggat aacgcgatca accagttgat cttgagggac 2100
cagatggaag cgctgggctg cagcgtggag ctgctcttcg atggtcgca ggcgttgctg 2160
cactgccaga cggcctgctt cgacgtgggt ctaccgata tcaacatgcc gaacatgaac 2220
ggatacgagc taaccgcgga gctacggcgc caagggttcc ggcagccgat catcggcgcg 2280
acggcgaacg ccatgcgtga ggagcgcgag cgctgcatgt ccgccgggat gaacgattgc 2340
ctggtcaaac cgggtggatct gaatgccctt cagaactgct tgattaatat tctcaagggtg 2400
gatcgatga 2409

```

Figure 6H

ORF1-8

SEQ ID NO:15

```

atgttgcgcc ggaatgatgc cggaaactac atgctgctgg atcagcatgg gcaggctgta 60
ctcgctacgg acgcagaggg gctggggagc ggtgcgtcgc ggacgctttt gcgtggagac 120
ggcttcggtt tcatcggtgc tggccactg ccgcagcata tgggtgcttt ccagcacgtg 180
gggtcttcga gctgggatct gatctatcac atcggtatcg gtcgcctgtt gctggctctg 240
tggctccctc tgttacttgc ctctgcgttg gcactgcag tcggcatcct actgcattgg 300
ctggtgcgga gcatcgagcg acgcttgata gagcccgcaa agcagcgcct tgaagcattg 360
aaggagagcg aagccttttc ccgtgcagtt atccaggccg cgcgcgtcgc gctgtgcgtg 420
ctgcgtcgtg ccgacgcgcg agtggtcctg gaaaatcccc aggcgcgcca atggctgggt 480
gatagcgagg cgattgcccga cgacgcgcgc agatggattt ccaggcggt cgaggagggt 540
gtgaagtgtt ctggagaaga actggaaacc gaggcagggc tacatcttca tctcaattac 600
acgcccaccc gctataacgg tgaagacgta ttgttctgcg cttcagtgga aatcagtgca 660
cgcaagcgga tggaggcgga actggctcgc gaaaaatccc tggcggtatg tgccaatgaa 720
gccaagacgc tgtttctcgc caccatgagc catgaaatcc gcacacctc gtacggcatg 780
ctggcacgc ttgagctgct tgggcgtacc gagctgagtc ggcagcaggc cggttacctg 840
aaggcaatcc agcatctctc gtcgacctg ctgcaactga tcagcgatgt gcttgacgta 900
tccaagatag aggcgcgcca actggacctg gactgcgtgg aattctcccc gctggaattg 960
accgaagagg tcgtgcagtc gttcaccggt gccgcgcagg ccaagggggt gcagttgtat 1020
acctgcctct ctgcggagct gccgctgcgc atgcgggggg ccgcggcgtc gatccggcag 1080
attctcaaca acctgctgag caacgcggtg aagttcacgc acaatggcta tgtcaacgtc 1140
cacctgaagg ccagcgtggg cgatgccgaa tgtgtgatgc tgacctggca ggtcaacgat 1200
accggcatgg ggatcaacgt cgaggatcag ccgcgtctgt tcgaaccgtt ctaccagata 1260
cgccgctccg agcatccggt cgaggcacg ggcctcggct tgcgatcag ccagcgctg 1320
gcgcagctaa tgaatggcag tctgaaactg gtcagtgagc tgggggtggg cagcagcttt 1380
agcctcaggc ttccgcttga gcggatcgcg atgcaggctg agccgcagga cctagccggg 1440
tgccgcgtcc aagtgtcggc gcctgtccgc gacctaacgg aatgcctgtg tggctggatc 1500
tcccgtggg gtggaagggc catggtcgcg acgccagggt cgctggacga ggcggacgcg 1560
acctcgctgc tggtcgaagt gttactgctg gagggggcgc cgatgttcga agcatggcca 1620
ggatgccggg tggagctttc cctcagggt gatatggagc cgcaggcaca gggccgcgac 1680
tggctgctcg ggctcaacaa cctggacggc ctgcatcgtg ctctgggctt ggcccattgg 1740
cgtctcgtcg atccttcgac gccgccgata cggctggctc cgttgcgcaa tctaggtctc 1800
cgcgtcctag tgggtggagga taacgcgata aaccagttga tcttgaggga ccagatggaa 1860
gcgctgggct gcagcgtgga gctgctcttc gatggtcgcg aggcgttgct gcactgccag 1920
acggcctgct tcgacgtggg gctcaccgat atcaacatgc cgaacatgaa cggatacgag 1980
ctaaccgcgg agctacggcg ccaagggttc cggcagccga tcatcggcgc gacggcgaac 2040
gccatgcgtg aggagcgcga gcgctgcag tccgcggga tgaacgattg cctggtaaaa 2100
ccggtggatc tgaatgccct tcagaactgc ttgattaata ttctcaagggt ggatcgatga 2160

```

Figure 6I

ORF1-9

SEQ ID NO:16

```

atgctgctgg atcagcatgg gcaggtcgta ctcgctacgg acgcagaggg gctggggagc 60
gggtgcgtgc ggacgctttt gcgtggagac ggcttcgggt tcatcggtgc tggcccaactg 120
ccgcagcata tgggtgctttt ccagcacgtg gggtcttcga gctgggatct gatctatcac 180
atcggtatcg gtcgcctgtt gctggctctg tggctccctc tgttacttgc ctctgcgttg 240
gcactcgcag tcggcatcct actgcattgg ctggtgcgga gcacgcagcg acgcttgata 300
gagcccgcaa agcgacgcct tgaagcattg aaggagagcg aagccttttc ccgtgcagtt 360
atccaggccg cgcccgtcgc gctgtgcgtg ctgctgcgtg ccgacgccgc agtggctctg 420
gaaaatcccc aggcgcgcca atggctgggt gatagcgagg cgattgcccc cgacgcgccg 480
agatggattt cccaggcggt cgcaggaggt gtgaagtgtt ctggagaaga actggaaacc 540
gaggcagggc tacatcttca tctcaattac acgcccaccc gctataacgg tgaagacgta 600
ttgttctgcg ccttcagtga aatcagtga cgcaagcgga tggaggcgga actggctcgc 660
gcaaaatccc tggcggtatg tgccaatgaa gccaaagacg tgtttctcgc caccatgagc 720
catgaaatcc gcacacctct gtacggcatg cttggcacgc ttgagctgct tgggcgtacc 780
gagctgagtc ggcagcaggc cggttaccta aaggcaatcc agcattcctc gtcgacctg 840
ctgcaactga tcagcgatgt gcttgacgta tccaagatag aggcgggcca actggacctg 900
gagtgcgtgg aattctcccc gctggaattg accgaagagg tcgtgcagtc gttcaccggt 960
gccgcgcagg ccaaggggct gcagttgtat acctgcctct ctgaggagct gccgctgcgc 1020
atgcgggggg ccgcggcgct gatccggcag attctcaaca acctgctgag caacgcggtg 1080
aagttcaccg acaatggcta tgtcaacgct cacctgaagg ccagcgtggt cgatgccgaa 1140
tgtgtgatgc tgacctggca ggtcaacgat accggcatgg ggatcaacgt cgaggatcag 1200
ccgcgtctgt tcgaaccgtt ctaccagata cgccgctccg agcatccggt cgcaggcacg 1260
ggcctcggtt tgtcgatcag ccagcgcctg gcgcagctaa tgaatggcag tctgaaactg 1320
gtcagtgagc tgggggttgg cagcagcttt agcctcaggc ttccgcttga gcggatcgcg 1380
atgcaggctg agccgcagga cctagccggg tgccgcgtcc aagtgcgtgg gcctgtccgc 1440
gacctaacgg aatgcctgtg tggctggatc tcccgtggtg gtggaagggc catggctcgc 1500
acgccgaggt cgctggacga ggcggacgcg acctcgtgc tggtcgaagt gttactgctg 1560
gagggggcgc cgatgttcga agcatggcca ggatgccggg tggagctttc cctcagggt 1620
gatatggagc cgcaggcaca gggccgcgac tggctgctcg ggctcaacaa cctggacggc 1680
ctgcatcgtg ctctgggcct ggcccatggg cgtctcgctg atccttcgac gccgcgcgata 1740
cggctggctc cgttgcgcaa tctaggtctc cgcgtcctag tgggtggagga taacgcgata 1800
aaccagttga tcttgaggga ccagatggaa gcgctgggct gcagcgtgga gctgctctc 1860
gatggctcgc aggcgttgct gcactgccag acggcctgct tcgacgtggt gctcaccgat 1920
atcaacatgc cgaacatgaa cggatacgag ctaaccgcgg agctacggcg ccaaggggtt 1980
cggcagccga tcatcggcgc gacggcgaaac gccatgcgtg aggagcgcga gcgctgcag 2040
tccgccggga tgaacgattg cctggtcaaa ccggtggatc tgaatgccct tcagaactgc 2100
ttgattaata ttctcaagggt ggatcgatga 2130

```


Figure 6J

ORF1-10
SEQ ID NO:17

```

atggtgcttt tccagcacgt ggggtcttcg agctgggacg tgatctatca catcggtatc 60
ggtcgccctgt tgctggctct gtggctccct ctgttacttg cctctgcgtt ggcactcgca 120
gtcggcatcc tactgcattg gctgggtgcg agcatcgagc gacgcttgat agagcccga 180
aagcgacgcc ttgaagcatt gaaggagagc gaagcctttt cccgtgcagt tatccaggcc 240
gcgcccgctg cgctgtgcgt gctgcgtcgt gccgacgccg cagtggctct ggaaaatccc 300
caggcgcgcc aatggctggg tgatagcgag gcgattgccc acgacgcgcc gagatggatt 360
tcccaggcgt tcgcaggagg tgtgaagtgt tctggagaag aactggaaac cgaggcaggg 420
ctacatcttc atctcaatta cagcccacc cgctataacg gtgaagacgt attgttctgc 480
gccttcagtg aaatcagtgc acgcaagcgg atggaggcgg aactggctcg cgaaaaatcc 540
ctggcggatg ctgccaatga agccaagacg ctgtttctcg ccaccatgag ccatgaaatc 600
cgcacacctc tgtacggcat gcttggcacg cttgagctgc ttgggcgtac cgagctgagt 660
cggcagcagg ccggttacct aaaggcaatc cagcattcct cgtcgaccct gctgcaactg 720
atcagcgatg tgcttgacgt atccaagata gaggccggcc aactggacct agagtgcgtg 780
gaattctccc cgctggaatt gaccgaagag gtcgtgcagt cgttcaccgg tgccgcgcag 840
gccaaggggc tgcagttgta tacctgcctc tctgcggagc tgccgctgcg catgcggggg 900
gccgcggcgt cgatccggca gattctcaac aacctgctga gcaacgcggg gaagttcacc 960
gacaatggct atgtcaacgt ccacctgaag gccagcgtgg tcgatgccga atgtgtgatg 1020
ctgacctggc aggtcaacga taccggcatg gggatcaacg tcgaggatca gccgcgtctg 1080
ttcgaaccgt tctaccagat acgccgctcc gagcatccgg tcgcaggcac gggcctcggc 1140
ttgtcgatca gccagcgctt tagcctcagg cttccgcttg agcggatcgc gatgcaggct 1200
ctgggggttg gcagcagctt tagcctcagg cttccgcttg agcggatcgc gatgcaggct 1260
gagccgcagg acctagccgg gtgcgcccgc caagtgcctg cgctgtccg cgacctaacg 1320
gaatgcctgt gtggctggat ctcccgtggt ggtggaaggg ccatggctcg gacgccgagg 1380
tcgctggacg aggcggacgc gacctcgctg ctggctgaag tggtactgct ggagggggcg 1440
ccgatgttcg aagcatggcc aggatgccgg gtggagcttt cccctcaggg tgatatggag 1500
ccgcaggcac agggccgcga ctggctgctc gggctcaaca acctggacgg cctgcatcgt 1560
gctctgggccc tggcccatgg gcgtctcgct gatccttcga cgccgccgat acggctggct 1620
ccgttgcgca atctaggtct ccgcgtccta gtggtggagg ataacgcgat caaccagttg 1680
atcttgaggg accagatgga agcgtcgggc tcgacgctgg agctgctctt cgatggctcg 1740
gaggcgttgc tgcaactgcca gacggcctgc ttcgacgtgg tgctcaccga tatcaacatg 1800
ccgaacatga acggatacga gctaaccgcg gagctacggc gccaaagggt ccggcgagccg 1860
atcatcggcg cgacggcgaa cgccatgcgt gaggagcgcg agcgtgcat gtccgcccgg 1920
atgaacgatt gcctggtcaa accggtggat ctgaatgccc ttcagaactg cttgattaat 1980
attctcaagg tggatcgatg a
2001

```

Figure 6K

ORF1-11

SEQ ID NO:18

```

atggaggcgg aactggctcg cgcaaaatcc ctggcggatg ctgccaatga agccaagacg 60
ctgtttctcg ccaccatgag ccatgaaatc cgcacacctc tgtacggcat gcttggcacg 120
cttgagctgc ttgggcgtac cgagctgagt cggcagcagg ccggttacct aaaggcaatc 180
cagcattcct cgtcgaccct gctgcaactg atcagcgatg tgcttgacgt atccaagata 240
gaggccggcc aactggacct agagtgcgtg gaattctccc cgctggaatt gaccgaagag 300
gtcgtgcagt cgttcacccg tgccgcgcag gccaaagggc tgcaattgta tacctgcctc 360
tctgcggagc tgccgctgcg catgcggggg gccgcggcgt cgatccggca gattctcaac 420
aacctgctga gcaacgcggg gaagttcacc gacaatggct atgtcaacgt ccacctgaag 480
gccagcgtgg tcgatgccga atgtgtgatg ctgacctggc aggtcaacga taccggcatg 540
gggatcaacg tcgaggatca gccgcgtctg ttcgaaccgt tctaccagat acgccgctcc 600
gagcatccgg tcgcaggcac gggcctcggc ttgtcgatca gccagcgcct ggcgcagcta 660
atgaatggca gtcgtaaaact ggtcagttag ctggggttgg gcagcagctt tagcctcagg 720
cttccgcttg agcggatcgc gatgcaggct gagccgcagg acctagccgg gtgcgccgtc 780
caagtgcctg cgctgtccg cgacctaacg gaatgcctgt gtggctggat ctcccgtgg 840
ggtaggaagg ccatggctgc gacgccgagg tcgctggacg aggcggacgc gacctcgctg 900
ctggtcgaag tgttactgct ggagggggcg ccgatgttcg aagcatggcc aggatgccgg 960
gtggagcttt ccctcagggg tgatatggag ccgcaggcac agggccgcga ctggctgctc 1020
gggctcaaca acctggacgg cctgcacgtg gctctggggc tggcccatgg gcgtctcgct 1080
gatccttcga cgccgccgat acggctggct ccgttgcgca atctaggtct ccgcgtccta 1140
gtggtggagg ataacgcgat caaccagttg atcttgaggg accagatgga agcgtgggc 1200
tgcagcgtgg agctgctctt cgatggtcgc gaggcgttgc tgcaactgca gacggcctgc 1260
ttcgacgtgg tgetcaccga tatcaacatg ccgaacatga acggatacga gctaaccgcg 1320
gagctacggc gccaaagggt ccggcagccg atcatcggcg cgacggcgaa cgccatgcgt 1380
gaggagcgcg agcgtgcat gtccgccggg atgaacgatt gcctgggtcaa accggtggat 1440
ctgaatgccc ttcagaactg cttgattaat attctcaagg tggatcgatg a 1491

```

Figure 6L

ORF1-1
 SEQ ID NO:19
 Met Lys Leu Lys Asn Phe Leu Gln Pro Phe Asp Ser Gly Phe Ser Thr
 1 5 10 15
 Pro Ser Ala Ala Leu Lys Leu Leu Arg Met Leu Gly Gly Ala Leu Met
 20 25 30
 Leu Cys Val Leu Cys Ser Leu Ile Phe Ser Val Ser Met Val Leu Asn
 35 40 45
 His Gln Val Ser Leu Ser Arg Gln Ala Met Asn Val Ala Met Tyr Glu
 50 55 60
 Ala Gln Leu Tyr Phe Glu Gln Arg Glu Ala Leu Leu Asn His Leu Ser
 65 70 75 80
 Gly Asn Val Val Pro Leu Ala Ala Gly Arg Ala Leu Val Asn Glu Ala
 85 90 95
 Pro Asn Asn Val Ser Ile Leu Pro Leu Ser Asp Gly Gly Arg Gly Leu
 100 105 110
 Leu Leu Thr Ala Arg Thr Leu Gly Asp Leu Arg Glu Lys Arg Leu Ala
 115 120 125
 Leu Met Tyr Leu Val Asp Thr Asp Lys Gly Pro Leu Val Tyr Arg Leu
 130 135 140
 Thr Ala Asp Gly Arg Pro Ser Ala Ala Ile Ser Ser Thr Ile Thr Lys
 145 150 155 160
 Glu Val Tyr Arg Ala Leu Leu Ala Thr Pro Ser Ala Pro Val His Trp
 165 170 175
 Val Thr Asp Gly Gly Thr Pro Gln Arg Leu Tyr Leu Phe Glu Ser Leu
 180 185 190
 Gly Asp Glu Pro Gly Glu Gly Trp Leu Gly Leu Glu Ile Leu Gly Glu
 195 200 205
 Asp Leu Asp Ser Met Leu Arg Arg Asn Asp Ala Gly Asn Tyr Met Leu
 210 215 220
 Leu Asp Gln His Gly Gln Val Val Leu Ala Thr Asp Ala Glu Ala Leu
 225 230 235 240
 Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp Gly Phe Gly Phe
 245 250 255
 Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu Phe Gln His Val
 260 265 270
 Gly Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly Ile Gly Arg Leu
 275 280 285
 Leu Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser Ala Leu Ala Leu
 290 295 300
 Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser Ile Glu Arg Arg
 305 310 315 320
 Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu Lys Glu Ser Glu
 325 330 335
 Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val Ala Leu Cys Val
 340 345 350
 Leu Arg Arg Ala Asp Ala Ala Val Val Leu Glu Asn Pro Gln Ala Arg
 355 360 365
 Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp Ala Pro Arg Trp
 370 375 380
 Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser Gly Glu Glu Leu
 385 390 395 400
 Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr Thr Pro Thr Arg
 405 410 415
 Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser Glu Ile Ser Ala
 420 425 430
 Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu Ala Asp
 435 440 445
 Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr Met Ser His Glu
 450 455 460

6L/1

Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu Leu Gly
 465 470 475 480
 Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala Ile Gln
 485 490 495
 His Ser Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp Val Leu Asp Val
 500 505 510
 Ser Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys Val Glu Phe Ser
 515 520 525
 Pro Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe Thr Gly Ala Ala
 530 535 540
 Gln Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser Ala Glu Leu Pro
 545 550 555 560
 Leu Arg Met Arg Gly Ala Ala Ala Ser Ile Arg Gln Ile Leu Asn Asn
 565 570 575
 Leu Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly Tyr Val Asn Val
 580 585 590
 His Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val Met Leu Thr Trp
 595 600 605
 Gln Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu Asp Gln Pro Arg
 610 615 620
 Leu Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu His Pro Val Ala
 625 630 635 640
 Gly Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln Leu Met
 645 650 655
 Asn Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu Gly Ser Ser Phe
 660 665 670
 Ser Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln Ala Glu Pro Gln
 675 680 685
 Asp Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro Val Arg Asp Leu
 690 695 700
 Thr Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg Ala Met
 705 710 715 720
 Val Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser Leu Leu
 725 730 735
 Val Glu Val Leu Leu Leu Glu Gly Ala Pro Met Phe Glu Ala Trp Pro
 740 745 750
 Gly Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met Glu Pro Gln Ala
 755 760 765
 Gln Gly Arg Asp Trp Leu Leu Gly Leu Asn Asn Leu Asp Gly Leu His
 770 775 780
 Arg Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp Pro Ser Thr Pro
 785 790 795 800
 Pro Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu Arg Val Leu Val
 805 810 815
 Val Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg Asp Gln Met Glu
 820 825 830
 Ala Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly Arg Glu Ala Leu
 835 840 845
 Leu His Cys Gln Thr Ala Cys Phe Asp Val Val Leu Thr Asp Ile Asn
 850 855 860
 Met Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu Leu Arg Arg Gln
 865 870 875 880
 Gly Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn Ala Met Arg Glu
 885 890 895
 Glu Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp Cys Leu Val Lys
 900 905 910
 Pro Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile Asn Ile Leu Lys
 915 920 925
 Val Asp Arg
 930

66/2

Figure 6M

ORF1-2

SEQ ID NO:20

Met Leu Gly Gly Ala Leu Met Leu Cys Val Leu Cys Ser Leu Ile Phe
 1 5 10 15
 Ser Val Ser Met Val Leu Asn His Gln Val Ser Leu Ser Arg Gln Ala
 20 25 30
 Met Asn Val Ala Met Tyr Glu Ala Gln Leu Tyr Phe Glu Gln Arg Glu
 35 40 45
 Ala Leu Leu Asn His Leu Ser Gly Asn Val Val Pro Leu Ala Ala Gly
 50 55 60
 Arg Ala Leu Val Asn Glu Ala Pro Asn Asn Val Ser Ile Leu Pro Leu
 65 70 75 80
 Ser Asp Gly Gly Arg Gly Leu Leu Leu Thr Ala Arg Thr Leu Gly Asp
 85 90 95
 Leu Arg Glu Lys Arg Leu Ala Leu Met Tyr Leu Val Asp Thr Asp Lys
 100 105 110
 Gly Pro Leu Val Tyr Arg Leu Thr Ala Asp Gly Arg Pro Ser Ala Ala
 115 120 125
 Ile Ser Ser Thr Ile Thr Lys Glu Val Tyr Arg Ala Leu Leu Ala Thr
 130 135 140
 Pro Ser Ala Pro Val His Trp Val Thr Asp Gly Gly Thr Pro Gln Arg
 145 150 155 160
 Leu Tyr Leu Phe Glu Ser Leu Gly Asp Glu Pro Gly Glu Gly Trp Leu
 165 170 175
 Gly Leu Glu Ile Leu Gly Glu Asp Leu Asp Ser Met Leu Arg Arg Asn
 180 185 190
 Asp Ala Gly Asn Tyr Met Leu Leu Asp Gln His Gly Gln Val Val Leu
 195 200 205
 Ala Thr Asp Ala Glu Ala Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu
 210 215 220
 Arg Gly Asp Gly Phe Gly Phe Ile Gly Ala Gly Pro Leu Pro Gln His
 225 230 235 240
 Met Val Leu Phe Gln His Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr
 245 250 255
 His Ile Gly Ile Gly Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu
 260 265 270
 Leu Ala Ser Ala Leu Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu
 275 280 285
 Val Arg Ser Ile Glu Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu
 290 295 300
 Glu Ala Leu Lys Glu Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala
 305 310 315 320
 Ala Pro Val Ala Leu Cys Val Leu Arg Arg Ala Asp Ala Val Val
 325 330 335
 Leu Glu Asn Pro Gln Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile
 340 345 350
 Ala His Asp Ala Pro Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val
 355 360 365
 Lys Cys Ser Gly Glu Glu Leu Glu Thr Glu Ala Gly Leu His Leu His
 370 375 380
 Leu Asn Tyr Thr Pro Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys
 385 390 395 400
 Ala Phe Ser Glu Ile Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala
 405 410 415
 Arg Ala Lys Ser Leu Ala Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe
 420 425 430
 Leu Ala Thr Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu
 435 440 445
 Gly Thr Leu Glu Leu Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala
 450 455 460

6 m/1

Gly Tyr Leu Lys Ala Ile Gln His Ser Ser Ser Thr Leu Leu Gln Leu
 465 470 475 480
 Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp
 485 490 495
 Leu Glu Cys Val Glu Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val
 500 505 510
 Gln Ser Phe Thr Gly Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr
 515 520 525
 Cys Leu Ser Ala Glu Leu Pro Leu Arg Met Arg Gly Ala Ala Ala Ser
 530 535 540
 Ile Arg Gln Ile Leu Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr
 545 550 555 560
 Asp Asn Gly Tyr Val Asn Val His Leu Lys Ala Ser Val Val Asp Ala
 565 570 575
 Glu Cys Val Met Leu Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile
 580 585 590
 Asn Val Glu Asp Gln Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg
 595 600 605
 Arg Ser Glu His Pro Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser
 610 615 620
 Gln Arg Leu Ala Gln Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu
 625 630 635 640
 Leu Gly Leu Gly Ser Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile
 645 650 655
 Ala Met Gln Ala Glu Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val
 660 665 670
 Leu Ala Pro Val Arg Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser
 675 680 685
 Arg Trp Gly Gly Arg Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu
 690 695 700
 Ala Asp Ala Thr Ser Leu Leu Val Glu Val Leu Leu Leu Glu Gly Ala
 705 710 715 720
 Pro Met Phe Glu Ala Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln
 725 730 735
 Gly Asp Met Glu Pro Gln Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu
 740 745 750
 Asn Asn Leu Asp Gly Leu His Arg Ala Leu Gly Leu Ala His Gly Arg
 755 760 765
 Leu Ala Asp Pro Ser Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn
 770 775 780
 Leu Gly Leu Arg Val Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu
 785 790 795 800
 Ile Leu Arg Asp Gln Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu
 805 810 815
 Phe Asp Gly Arg Glu Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp
 820 825 830
 Val Val Leu Thr Asp Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu
 835 840 845
 Thr Ala Glu Leu Arg Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala
 850 855 860
 Thr Ala Asn Ala Met Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly
 865 870 875 880
 Met Asn Asp Cys Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn
 885 890 895
 Cys Leu Ile Asn Ile Leu Lys Val Asp Arg
 900 905

6m/2

Figure 6N

SEQ ID NO:21

ORF1-3

Met Leu Cys Val Leu Cys Ser Leu Ile Phe Ser Val Ser Met Val Leu
 1 5 10 15
 Asn His Gln Val Ser Leu Ser Arg Gln Ala Met Asn Val Ala Met Tyr
 20 25 30
 Glu Ala Gln Val Tyr Phe Glu Gln Arg Glu Ala Leu Leu Asn His Leu
 35 40 45
 Ser Gly Asn Val Val Pro Leu Ala Ala Gly Arg Ala Leu Val Asn Glu
 50 55 60
 Ala Pro Asn Asn Val Ser Ile Leu Pro Leu Ser Asp Gly Gly Arg Gly
 65 70 75 80
 Leu Leu Leu Thr Ala Arg Thr Leu Gly Asp Leu Arg Glu Lys Arg Leu
 85 90 95
 Ala Leu Met Tyr Leu Val Asp Thr Asp Lys Gly Pro Leu Val Tyr Arg
 100 105 110
 Leu Thr Ala Asp Gly Arg Pro Ser Ala Ala Ile Ser Ser Thr Ile Thr
 115 120 125
 Lys Glu Val Tyr Arg Ala Leu Ala Thr Pro Ser Ala Pro Val His
 130 135 140
 Trp Val Thr Asp Gly Gly Thr Pro Gln Arg Leu Tyr Leu Phe Glu Ser
 145 150 155 160
 Leu Gly Asp Glu Pro Gly Glu Gly Trp Leu Gly Leu Glu Ile Leu Gly
 165 170 175
 Glu Asp Leu Asp Ser Met Leu Arg Arg Asn Asp Ala Gly Asn Tyr Met
 180 185 190
 Leu Leu Asp Gln His Gly Gln Val Val Leu Ala Thr Asp Ala Glu Ala
 195 200 205
 Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp Gly Phe Gly
 210 215 220
 Phe Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu Phe Gln His
 225 230 235 240
 Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly Ile Gly Arg
 245 250 255
 Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser Ala Leu Ala
 260 265 270
 Leu Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser Ile Glu Arg
 275 280 285
 Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu Lys Glu Ser
 290 295 300
 Glu Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val Ala Leu Cys
 305 310 315 320
 Val Leu Arg Arg Ala Asp Ala Ala Val Val Leu Glu Asn Pro Gln Ala
 325 330 335
 Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp Ala Pro Arg
 340 345 350
 Trp Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser Gly Glu Glu
 355 360 365
 Leu Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr Thr Pro Thr
 370 375 380
 Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser Glu Ile Ser
 385 390 395 400
 Ala Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu Ala
 405 410 415
 Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr Met Ser His
 420 425 430
 Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu Leu
 435 440 445
 Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala Ile
 450 455 460

6N/1

Gln	His	Ser	Ser	Ser	Thr	Leu	Leu	Gln	Leu	Ile	Ser	Asp	Val	Leu	Asp
465					470					475					480
Val	Ser	Lys	Ile	Glu	Ala	Gly	Gln	Leu	Asp	Leu	Glu	Cys	Val	Glu	Phe
				485					490						495
Ser	Pro	Leu	Glu	Leu	Thr	Glu	Glu	Val	Val	Gln	Ser	Phe	Thr	Gly	Ala
			500					505					510		
Ala	Gln	Ala	Lys	Gly	Leu	Gln	Leu	Tyr	Thr	Cys	Leu	Ser	Ala	Glu	Leu
		515					520					525			
Pro	Leu	Arg	Met	Arg	Gly	Ala	Ala	Ala	Ser	Ile	Arg	Gln	Ile	Leu	Asn
		530				535					540				
Asn	Leu	Leu	Ser	Asn	Ala	Val	Lys	Phe	Thr	Asp	Asn	Gly	Tyr	Val	Asn
545				550						555					560
Val	His	Leu	Lys	Ala	Ser	Val	Val	Asp	Ala	Glu	Cys	Val	Met	Leu	Thr
				565					570						575
Trp	Gln	Val	Asn	Asp	Thr	Gly	Met	Gly	Ile	Asn	Val	Glu	Asp	Gln	Pro
			580					585					590		
Arg	Leu	Phe	Glu	Pro	Phe	Tyr	Gln	Ile	Arg	Arg	Ser	Glu	His	Pro	Val
		595					600					605			
Ala	Gly	Thr	Gly	Leu	Gly	Leu	Ser	Ile	Ser	Gln	Arg	Leu	Ala	Gln	Leu
		610				615					620				
Met	Asn	Gly	Ser	Leu	Lys	Leu	Val	Ser	Glu	Leu	Gly	Leu	Gly	Ser	Ser
625					630					635					640
Phe	Ser	Leu	Arg	Leu	Pro	Leu	Glu	Arg	Ile	Ala	Met	Gln	Ala	Glu	Pro
				645					650						655
Gln	Asp	Leu	Ala	Gly	Cys	Ala	Val	Gln	Val	Leu	Ala	Pro	Val	Arg	Asp
			660					665					670		
Leu	Thr	Glu	Cys	Leu	Cys	Gly	Trp	Ile	Ser	Arg	Trp	Gly	Gly	Arg	Ala
		675					680					685			
Met	Val	Ala	Thr	Pro	Arg	Ser	Leu	Asp	Glu	Ala	Asp	Ala	Thr	Ser	Leu
		690				695					700				
Leu	Val	Glu	Val	Leu	Leu	Glu	Gly	Ala	Pro	Met	Phe	Glu	Ala	Trp	
705					710					715					720
Pro	Gly	Cys	Arg	Val	Glu	Leu	Ser	Pro	Gln	Gly	Asp	Met	Glu	Pro	Gln
				725					730						735
Ala	Gln	Gly	Arg	Asp	Trp	Leu	Leu	Gly	Leu	Asn	Asn	Leu	Asp	Gly	Leu
			740					745					750		
His	Arg	Ala	Leu	Gly	Leu	Ala	His	Gly	Arg	Leu	Ala	Asp	Pro	Ser	Thr
		755					760					765			
Pro	Pro	Ile	Arg	Leu	Ala	Pro	Leu	Arg	Asn	Leu	Gly	Leu	Arg	Val	Leu
		770				775					780				
Val	Val	Glu	Asp	Asn	Ala	Ile	Asn	Gln	Leu	Ile	Leu	Arg	Asp	Gln	Met
785					790					795					800
Glu	Ala	Leu	Gly	Cys	Ser	Val	Glu	Leu	Leu	Phe	Asp	Gly	Arg	Glu	Ala
				805					810						815
Leu	Leu	His	Cys	Gln	Thr	Ala	Cys	Phe	Asp	Val	Val	Leu	Thr	Asp	Ile
			820					825					830		
Asn	Met	Pro	Asn	Met	Asn	Gly	Tyr	Glu	Leu	Thr	Ala	Glu	Leu	Arg	Arg
		835					840					845			
Gln	Gly	Phe	Arg	Gln	Pro	Ile	Ile	Gly	Ala	Thr	Ala	Asn	Ala	Met	Arg
		850				855					860				
Glu	Glu	Arg	Glu	Arg	Cys	Met	Ser	Ala	Gly	Met	Asn	Asp	Cys	Leu	Val
865					870					875					880
Lys	Pro	Val	Asp	Leu	Asn	Ala	Leu	Gln	Asn	Cys	Leu	Ile	Asn	Ile	Leu
				885					890						895
Lys	Val	Asp	Arg												
			900												

6N/2

Figure 60

ORF1-4

SEQ ID NO:22

```

Met Val Leu Asn His Gln Val Ser Leu Ser Arg Gln Ala Met Asn Val
 1          5          10          15
Ala Met Tyr Glu Ala Gln Leu Tyr Phe Glu Gln Arg Glu Ala Leu Leu
 20          25          30
Asn His Leu Ser Gly Asn Val Val Pro Leu Ala Ala Gly Arg Ala Leu
 35          40          45
Val Asn Glu Ala Pro Asn Asn Val Ser Ile Leu Pro Leu Ser Asp Gly
 50          55          60
Gly Arg Gly Leu Leu Leu Thr Ala Arg Thr Leu Gly Asp Leu Arg Glu
 65          70          75          80
Lys Arg Leu Ala Leu Met Tyr Leu Val Asp Thr Asp Lys Gly Pro Leu
 85          90          95
Val Tyr Arg Leu Thr Ala Asp Gly Arg Pro Ser Ala Ala Ile Ser Ser
 100         105         110
Thr Ile Thr Lys Glu Val Tyr Arg Ala Leu Leu Ala Thr Pro Ser Ala
 115         120         125
Pro Val His Trp Val Thr Asp Gly Gly Thr Pro Gln Arg Leu Tyr Leu
 130         135         140
Phe Glu Ser Leu Gly Asp Glu Pro Gly Glu Gly Trp Leu Gly Leu Glu
 145         150         155         160
Ile Leu Gly Glu Asp Leu Asp Ser Met Leu Arg Arg Asn Asp Ala Gly
 165         170         175
Asn Tyr Met Leu Leu Asp Gln His Gly Gln Val Val Leu Ala Thr Asp
 180         185         190
Ala Glu Ala Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp
 195         200         205
Gly Phe Gly Phe Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu
 210         215         220
Phe Gln His Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly
 225         230         235         240
Ile Gly Arg Leu Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser
 245         250         255
Ala Leu Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser
 260         265         270
Ile Glu Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu
 275         280         285
Lys Glu Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val
 290         295         300
Ala Leu Cys Val Leu Arg Arg Ala Asp Ala Val Val Leu Glu Asn
 305         310         315         320
Pro Gln Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp
 325         330         335
Ala Pro Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser
 340         345         350
Gly Glu Glu Leu Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr
 355         360         365
Thr Pro Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser
 370         375         380
Glu Ile Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys
 385         390         395         400
Ser Leu Ala Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr
 405         410         415
Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu
 420         425         430
Glu Leu Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu
 435         440         445
Lys Ala Ile Gln His Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp
 450         455         460

```

60/1

Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys
 465 470 475 480
 Val Glu Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe
 485 490 495
 Thr Gly Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser
 500 505 510
 Ala Glu Leu Pro Leu Arg Met Arg Gly Ala Ala Ala Ser Ile Arg Gln
 515 520 525
 Ile Leu Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly
 530 535 540
 Tyr Val Asn Val His Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val
 545 550 555 560
 Met Leu Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu
 565 570 575
 Asp Gln Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu
 580 585 590
 His Pro Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu
 595 600 605
 Ala Gln Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu
 610 615 620
 Gly Ser Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln
 625 630 635 640
 Ala Glu Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro
 645 650 655
 Val Arg Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly
 660 665 670
 Gly Arg Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala
 675 680 685
 Thr Ser Leu Leu Val Glu Val Leu Leu Glu Gly Ala Pro Met Phe
 690 695 700
 Glu Ala Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met
 705 710 715 720
 Glu Pro Gln Ala Gln Gly Arg Asp Trp Leu Gly Leu Asn Asn Leu
 725 730 735
 Asp Gly Leu His Arg Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp
 740 745 750
 Pro Ser Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu
 755 760 765
 Arg Val Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg
 770 775 780
 Asp Gln Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly
 785 790 795 800
 Arg Glu Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp Val Val Leu
 805 810 815
 Thr Asp Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu
 820 825 830
 Leu Arg Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn
 835 840 845
 Ala Met Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp
 850 855 860
 Cys Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile
 865 870 875 880
 Asn Ile Leu Lys Val Asp Arg
 885

60/2

Figure 6P

ORF1-5

SEQ ID NO:23

Met	Asn	Val	Ala	Met	Tyr	Glu	Ala	Gln	Leu	Tyr	Phe	Glu	Gln	Arg	Glu
1				5				10						15	
Ala	Leu	Leu	Asn	His	Leu	Ser	Gly	Asn	Val	Val	Pro	Leu	Ala	Ala	Gly
			20					25					30		
Arg	Ala	Leu	Val	Asn	Glu	Ala	Pro	Asn	Asn	Val	Ser	Ile	Leu	Pro	Leu
		35					40					45			
Ser	Asp	Gly	Gly	Arg	Gly	Leu	Leu	Leu	Thr	Ala	Arg	Thr	Leu	Gly	Asp
	50					55					60				
Leu	Arg	Glu	Lys	Arg	Leu	Ala	Leu	Met	Tyr	Leu	Val	Asp	Thr	Asp	Lys
65					70					75				80	
Gly	Pro	Leu	Val	Tyr	Arg	Leu	Thr	Ala	Asp	Gly	Arg	Pro	Ser	Ala	Ala
				85					90					95	
Ile	Ser	Ser	Thr	Ile	Thr	Lys	Glu	Val	Tyr	Arg	Ala	Leu	Leu	Ala	Thr
			100					105						110	
Pro	Ser	Ala	Pro	Val	His	Trp	Val	Thr	Asp	Gly	Gly	Thr	Pro	Gln	Arg
		115					120					125			
Leu	Tyr	Leu	Phe	Glu	Ser	Leu	Gly	Asp	Glu	Pro	Gly	Glu	Gly	Trp	Leu
	130						135				140				
Gly	Leu	Glu	Ile	Leu	Gly	Glu	Asp	Leu	Asp	Ser	Met	Leu	Arg	Arg	Asn
145					150					155				160	
Asp	Ala	Gly	Asn	Tyr	Met	Leu	Leu	Asp	Gln	His	Gly	Gln	Val	Val	Leu
			165						170					175	
Ala	Thr	Asp	Ala	Glu	Ala	Leu	Gly	Ser	Gly	Ala	Ser	Arg	Thr	Leu	Leu
		180						185					190		
Arg	Gly	Asp	Gly	Phe	Gly	Phe	Ile	Gly	Ala	Gly	Pro	Leu	Pro	Gln	His
		195					200					205			
Met	Val	Leu	Phe	Gln	His	Val	Gly	Ser	Ser	Ser	Trp	Asp	Leu	Ile	Tyr
	210					215					220				
His	Ile	Gly	Ile	Gly	Arg	Leu	Leu	Leu	Ala	Leu	Trp	Leu	Pro	Leu	Leu
225					230					235				240	
Leu	Ala	Ser	Ala	Leu	Ala	Leu	Ala	Val	Gly	Ile	Leu	Leu	His	Trp	Leu
			245						250					255	
Val	Arg	Ser	Ile	Glu	Arg	Arg	Leu	Ile	Glu	Pro	Ala	Lys	Arg	Arg	Leu
		260						265				270			
Glu	Ala	Leu	Lys	Glu	Ser	Glu	Ala	Phe	Ser	Arg	Ala	Val	Ile	Gln	Ala
		275					280					285			
Ala	Pro	Val	Ala	Leu	Cys	Val	Leu	Arg	Arg	Ala	Asp	Ala	Ala	Val	Val
	290					295					300				
Leu	Glu	Asn	Pro	Gln	Ala	Arg	Gln	Trp	Leu	Gly	Asp	Ser	Glu	Ala	Ile
305					310					315				320	
Ala	His	Asp	Ala	Pro	Arg	Trp	Ile	Ser	Gln	Ala	Phe	Ala	Gly	Gly	Val
			325						330					335	
Lys	Cys	Ser	Gly	Glu	Glu	Leu	Glu	Thr	Glu	Ala	Gly	Leu	His	Leu	His
		340						345					350		
Leu	Asn	Tyr	Thr	Pro	Thr	Arg	Tyr	Asn	Gly	Glu	Asp	Val	Leu	Phe	Cys
		355					360					365			
Ala	Phe	Ser	Glu	Ile	Ser	Ala	Arg	Lys	Arg	Met	Glu	Ala	Glu	Leu	Ala
	370					375					380				
Arg	Ala	Lys	Ser	Leu	Ala	Asp	Ala	Ala	Asn	Glu	Ala	Lys	Thr	Leu	Phe
385					390					395				400	
Leu	Ala	Thr	Met	Ser	His	Glu	Ile	Arg	Thr	Pro	Leu	Tyr	Gly	Met	Leu
			405						410					415	
Gly	Thr	Leu	Glu	Leu	Leu	Gly	Arg	Thr	Glu	Leu	Ser	Arg	Gln	Gln	Ala
		420						425					430		
Gly	Tyr	Leu	Lys	Ala	Ile	Gln	His	Ser	Ser	Ser	Thr	Leu	Leu	Gln	Leu
	435						440					445			
Ile	Ser	Asp	Val	Leu	Asp	Val	Ser	Lys	Ile	Glu	Ala	Gly	Gln	Leu	Asp
	450					455					460				

6P/1

Leu Glu Cys Val Glu Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val
 465 470 475 480
 Gln Ser Phe Thr Gly Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr
 485 490 495
 Cys Leu Ser Ala Glu Leu Pro Leu Arg Met Arg Gly Ala Ala Ser
 500 505 510
 Ile Arg Gln Ile Leu Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr
 515 520 525
 Asp Asn Gly Tyr Val Asn Val His Leu Lys Ala Ser Val Val Asp Ala
 530 535 540
 Glu Cys Val Met Leu Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile
 545 550 555 560
 Asn Val Glu Asp Gln Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg
 565 570 575
 Arg Ser Glu His Pro Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser
 580 585 590
 Gln Arg Leu Ala Gln Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu
 595 600 605
 Leu Gly Leu Gly Ser Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile
 610 615 620
 Ala Met Gln Ala Glu Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val
 625 630 635 640
 Leu Ala Pro Val Arg Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser
 645 650 655
 Arg Trp Gly Gly Arg Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu
 660 665 670
 Ala Asp Ala Thr Ser Leu Leu Val Glu Val Leu Leu Leu Glu Gly Ala
 675 680 685
 Pro Met Phe Glu Ala Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln
 690 695 700
 Gly Asp Met Glu Pro Gln Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu
 705 710 715 720
 Asn Asn Leu Asp Gly Leu His Arg Ala Leu Gly Leu Ala His Gly Arg
 725 730 735
 Leu Ala Asp Pro Ser Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn
 740 745 750
 Leu Gly Leu Arg Val Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu
 755 760 765
 Ile Leu Arg Asp Gln Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu
 770 775 780
 Phe Asp Gly Arg Glu Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp
 785 790 795 800
 Val Val Leu Thr Asp Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu
 805 810 815
 Thr Ala Glu Leu Arg Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala
 820 825 830
 Thr Ala Asn Ala Met Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly
 835 840 845
 Met Asn Asp Cys Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn
 850 855 860
 Cys Leu Ile Asn Ile Leu Lys Val Asp Arg
 865 870

6P/2

Figure 6Q

ORF1-6
 SEQ ID NO:24
 Met Tyr Glu Ala Gln Leu Tyr Phe Glu Gln Arg Glu Ala Leu Leu Asn
 1 5 10 15
 His Leu Ser Gly Asn Val Val Pro Leu Ala Ala Gly Arg Ala Leu Val
 20 25 30
 Asn Glu Ala Pro Asn Asn Val Ser Ile Leu Pro Leu Ser Asp Gly Gly
 35 40 45
 Arg Gly Leu Leu Leu Thr Ala Arg Thr Leu Gly Asp Leu Arg Glu Lys
 50 55 60
 Arg Leu Ala Leu Met Tyr Leu Val Asp Thr Asp Lys Gly Pro Leu Val
 65 70 75 80
 Tyr Arg Leu Thr Ala Asp Gly Arg Pro Ser Ala Ala Ile Ser Ser Thr
 85 90 95
 Ile Thr Lys Glu Val Tyr Arg Ala Leu Leu Ala Thr Pro Ser Ala Pro
 100 105 110
 Val His Trp Val Thr Asp Gly Gly Thr Pro Gln Arg Leu Tyr Leu Phe
 115 120 125
 Glu Ser Leu Gly Asp Glu Pro Gly Glu Gly Trp Leu Gly Leu Glu Ile
 130 135 140
 Leu Gly Glu Asp Leu Asp Ser Met Leu Arg Arg Asn Asp Ala Gly Asn
 145 150 155 160
 Tyr Met Leu Leu Asp Gln His Gly Gln Val Val Leu Ala Thr Asp Ala
 165 170 175
 Glu Ala Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp Gly
 180 185 190
 Phe Gly Phe Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu Phe
 195 200 205
 Gln His Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly Ile
 210 215 220
 Gly Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser Ala
 225 230 235 240
 Leu Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser Ile
 245 250 255
 Glu Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu Lys
 260 265 270
 Glu Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val Ala
 275 280 285
 Leu Cys Val Leu Arg Arg Ala Asp Ala Ala Val Val Leu Glu Asn Pro
 290 295 300
 Gln Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp Ala
 305 310 315 320
 Pro Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser Gly
 325 330 335
 Glu Glu Leu Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr Thr
 340 345 350
 Pro Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser Glu
 355 360 365
 Ile Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys Ser
 370 375 380
 Leu Ala Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr Met
 385 390 395 400
 Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu
 405 410 415
 Leu Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys
 420 425 430
 Ala Ile Gln His Ser Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp Val
 435 440 445
 Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys Val
 450 455 460

6a/1

Glu Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe Thr
 465 470 475 480
 Gly Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser Ala
 485 490 495
 Glu Leu Pro Leu Arg Met Arg Gly Ala Ala Ala Ser Ile Arg Gln Ile
 500 505 510
 Leu Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly Tyr
 515 520 525
 Val Asn Val His Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val Met
 530 535 540
 Leu Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu Asp
 545 550 555 560
 Gln Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu His
 565 570 575
 Pro Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala
 580 585 590
 Gln Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu Gly
 595 600 605
 Ser Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln Ala
 610 615 620
 Glu Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro Val
 625 630 635 640
 Arg Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly
 645 650 655
 Arg Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr
 660 665 670
 Ser Leu Leu Val Glu Val Leu Leu Leu Glu Gly Ala Pro Met Phe Glu
 675 680 685
 Ala Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met Glu
 690 695 700
 Pro Gln Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu Asn Asn Leu Asp
 705 710 715 720
 Gly Leu His Arg Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp Pro
 725 730 735
 Ser Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu Arg
 740 745 750
 Val Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg Asp
 755 760 765
 Gln Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly Arg
 770 775 780
 Glu Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp Val Val Leu Thr
 785 790 795 800
 Asp Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu Leu
 805 810 815
 Arg Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn Ala
 820 825 830
 Met Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp Cys
 835 840 845
 Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile Asn
 850 855 860
 Ile Leu Lys Val Asp Arg
 865 870

6Q/2

Figure 6R

ORF1-7

SEQ ID NO:25

Met	Tyr	Leu	Val	Asp	Thr	Asp	Lys	Gly	Pro	Leu	Val	Tyr	Arg	Leu	Thr
1				5				10						15	
Ala	Asp	Gly	Arg	Pro	Ser	Ala	Ala	Ile	Ser	Ser	Thr	Ile	Thr	Lys	Glu
		20						25				30			
Val	Tyr	Arg	Ala	Leu	Leu	Ala	Thr	Pro	Ser	Ala	Pro	Val	His	Trp	Val
	35						40				45				
Thr	Asp	Gly	Gly	Thr	Pro	Gln	Arg	Leu	Tyr	Leu	Phe	Glu	Ser	Leu	Gly
	50					55					60				
Asp	Glu	Pro	Gly	Glu	Gly	Trp	Leu	Gly	Leu	Glu	Ile	Leu	Gly	Glu	Asp
65					70					75				80	
Leu	Asp	Ser	Met	Leu	Arg	Arg	Asn	Asp	Ala	Gly	Asn	Tyr	Met	Leu	Leu
			85						90					95	
Asp	Gln	His	Gly	Gln	Val	Val	Leu	Ala	Thr	Asp	Ala	Glu	Ala	Leu	Gly
			100					105					110		
Ser	Gly	Ala	Ser	Arg	Thr	Leu	Leu	Arg	Gly	Asp	Gly	Phe	Gly	Phe	Ile
	115						120					125			
Gly	Ala	Gly	Pro	Leu	Pro	Gln	His	Met	Val	Leu	Phe	Gln	His	Val	Gly
	130					135					140				
Ser	Ser	Ser	Trp	Asp	Leu	Ile	Tyr	His	Ile	Gly	Ile	Gly	Arg	Leu	Leu
145					150					155				160	
Leu	Ala	Leu	Trp	Leu	Pro	Leu	Leu	Leu	Ala	Ser	Ala	Leu	Ala	Leu	Ala
			165						170					175	
Val	Gly	Ile	Leu	Leu	His	Trp	Leu	Val	Arg	Ser	Ile	Glu	Arg	Arg	Leu
		180						185					190		
Ile	Glu	Pro	Ala	Lys	Arg	Arg	Leu	Glu	Ala	Leu	Lys	Glu	Ser	Glu	Ala
	195						200					205			
Phe	Ser	Arg	Ala	Val	Ile	Gln	Ala	Ala	Pro	Val	Ala	Leu	Cys	Val	Leu
	210					215					220				
Arg	Arg	Ala	Asp	Ala	Ala	Val	Val	Leu	Glu	Asn	Pro	Gln	Ala	Arg	Gln
225					230					235				240	
Trp	Leu	Gly	Asp	Ser	Glu	Ala	Ile	Ala	His	Asp	Ala	Pro	Arg	Trp	Ile
			245						250					255	
Ser	Gln	Ala	Phe	Ala	Gly	Gly	Val	Lys	Cys	Ser	Gly	Glu	Glu	Leu	Glu
		260						265					270		
Thr	Glu	Ala	Gly	Leu	His	Leu	His	Leu	Asn	Tyr	Thr	Pro	Thr	Arg	Tyr
	275						280					285			
Asn	Gly	Glu	Asp	Val	Leu	Phe	Cys	Ala	Phe	Ser	Glu	Ile	Ser	Ala	Arg
	290					295					300				
Lys	Arg	Met	Glu	Ala	Glu	Leu	Ala	Arg	Ala	Lys	Ser	Leu	Ala	Asp	Ala
305					310					315				320	
Ala	Asn	Glu	Ala	Lys	Thr	Leu	Phe	Leu	Ala	Thr	Met	Ser	His	Glu	Ile
			325						330					335	
Arg	Thr	Pro	Leu	Tyr	Gly	Met	Leu	Gly	Thr	Leu	Glu	Leu	Leu	Gly	Arg
		340						345					350		
Thr	Glu	Leu	Ser	Arg	Gln	Gln	Ala	Gly	Tyr	Leu	Lys	Ala	Ile	Gln	His
	355						360					365			
Ser	Ser	Ser	Thr	Leu	Leu	Gln	Leu	Ile	Ser	Asp	Val	Leu	Asp	Val	Ser
	370					375					380				
Lys	Ile	Glu	Ala	Gly	Gln	Leu	Asp	Leu	Glu	Cys	Val	Glu	Phe	Ser	Pro
385					390					395				400	
Leu	Glu	Leu	Thr	Glu	Glu	Val	Val	Gln	Ser	Phe	Thr	Gly	Ala	Ala	Gln
			405						410					415	
Ala	Lys	Gly	Leu	Gln	Leu	Tyr	Thr	Cys	Leu	Ser	Ala	Glu	Leu	Pro	Leu
		420						425					430		
Arg	Met	Arg	Gly	Ala	Ala	Ala	Ser	Ile	Arg	Gln	Ile	Leu	Asn	Asn	Leu
	435						440					445			
Leu	Ser	Asn	Ala	Val	Lys	Phe	Thr	Asp	Asn	Gly	Tyr	Val	Asn	Val	His
	450					455					460				

6R/1

Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val Met Leu Thr Trp Gln
 465 470 475 480
 Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu Asp Gln Pro Arg Leu
 485 490 495
 Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu His Pro Val Ala Gly
 500 505 510
 Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln Leu Met Asn
 515 520 525
 Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu Gly Ser Ser Phe Ser
 530 535 540
 Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln Ala Glu Pro Gln Asp
 545 550 555 560
 Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro Val Arg Asp Leu Thr
 565 570 575
 Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg Ala Met Val
 580 585 590
 Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser Leu Leu Val
 595 600 605
 Glu Val Leu Leu Leu Glu Gly Ala Pro Met Phe Glu Ala Trp Pro Gly
 610 615 620
 Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met Glu Pro Gln Ala Gln
 625 630 635 640
 Gly Arg Asp Trp Leu Leu Gly Leu Asn Asn Leu Asp Gly Leu His Arg
 645 650 655
 Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp Pro Ser Thr Pro Pro
 660 665 670
 Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu Arg Val Leu Val Val
 675 680 685
 Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg Asp Gln Met Glu Ala
 690 695 700
 Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly Arg Glu Ala Leu Leu
 705 710 715 720
 His Cys Gln Thr Ala Cys Phe Asp Val Val Leu Thr Asp Ile Asn Met
 725 730 735
 Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu Leu Arg Arg Gln Gly
 740 745 750
 Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn Ala Met Arg Glu Glu
 755 760 765
 Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp Cys Leu Val Lys Pro
 770 775 780
 Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile Asn Ile Leu Lys Val
 785 790 795 800
 Asp Arg

6R/2

Figure 6S

ORF1-8

SEQ ID NO:26

```

Met Leu Arg Arg Asn Asp Ala Gly Asn Tyr Met Leu Leu Asp Gln His
1      5      10      15
Gly Gln Val Val Leu Ala Thr Asp Ala Glu Ala Leu Gly Ser Gly Ala
20      25      30
Ser Arg Thr Leu Leu Arg Gly Asp Gly Phe Gly Phe Ile Gly Ala Gly
35      40      45
Pro Leu Pro Gln His Met Val Leu Phe Gln His Val Gly Ser Ser Ser
50      55      60
Trp Asp Leu Ile Tyr His Ile Gly Ile Gly Arg Leu Leu Leu Ala Leu
65      70      75      80
Trp Leu Pro Leu Leu Leu Ala Ser Ala Leu Ala Leu Ala Val Gly Ile
85      90      95
Leu Leu His Trp Leu Val Arg Ser Ile Glu Arg Arg Leu Ile Glu Pro
100     105     110
Ala Lys Arg Arg Leu Glu Ala Leu Lys Glu Ser Glu Ala Phe Ser Arg
115     120     125
Ala Val Ile Gln Ala Ala Pro Val Ala Leu Cys Val Leu Arg Arg Ala
130     135     140
Asp Ala Ala Val Val Leu Glu Asn Pro Gln Ala Arg Gln Trp Leu Gly
145     150     155     160
Asp Ser Glu Ala Ile Ala His Asp Ala Pro Arg Trp Ile Ser Gln Ala
165     170     175
Phe Ala Gly Gly Val Lys Cys Ser Gly Glu Glu Leu Glu Thr Glu Ala
180     185     190
Gly Leu His Leu His Leu Asn Tyr Thr Pro Thr Arg Tyr Asn Gly Glu
195     200     205
Asp Val Leu Phe Cys Ala Phe Ser Glu Ile Ser Ala Arg Lys Arg Met
210     215     220
Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu Ala Asp Ala Ala Asn Glu
225     230     235     240
Ala Lys Thr Leu Phe Leu Ala Thr Met Ser His Glu Ile Arg Thr Pro
245     250     255
Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu Leu Gly Arg Thr Glu Leu
260     265     270
Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala Ile Gln His Ser Ser Ser
275     280     285
Thr Leu Leu Gln Leu Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu
290     295     300
Ala Gly Gln Leu Asp Leu Glu Cys Val Glu Phe Ser Pro Leu Glu Leu
305     310     315     320
Thr Glu Glu Val Val Gln Ser Phe Thr Gly Ala Ala Gln Ala Lys Gly
325     330     335
Leu Gln Leu Tyr Thr Cys Leu Ser Ala Glu Leu Pro Leu Arg Met Arg
340     345     350
Gly Ala Ala Ala Ser Ile Arg Gln Ile Leu Asn Asn Leu Leu Ser Asn
355     360     365
Ala Val Lys Phe Thr Asp Asn Gly Tyr Val Asn Val His Leu Lys Ala
370     375     380
Ser Val Val Asp Ala Glu Cys Val Met Leu Thr Trp Gln Val Asn Asp
385     390     395     400
Thr Gly Met Gly Ile Asn Val Glu Asp Gln Pro Arg Leu Phe Glu Pro
405     410     415
Phe Tyr Gln Ile Arg Arg Ser Glu His Pro Val Ala Gly Thr Gly Leu
420     425     430
Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln Leu Met Asn Gly Ser Leu
435     440     445
Lys Leu Val Ser Glu Leu Gly Leu Gly Ser Ser Phe Ser Leu Arg Leu
450     455     460

```

6S/1

Pro Leu Glu Arg Ile Ala Met Gln Ala Glu Pro Gln Asp Leu Ala Gly
 465 470 475 480
 Cys Ala Val Gln Val Leu Ala Pro Val Arg Asp Leu Thr Glu Cys Leu
 485 490 495
 Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg Ala Met Val Ala Thr Pro
 500 505 510
 Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser Leu Leu Val Glu Val Leu
 515 520 525
 Leu Leu Glu Gly Ala Pro Met Phe Glu Ala Trp Pro Gly Cys Arg Val
 530 535 540
 Glu Leu Ser Pro Gln Gly Asp Met Glu Pro Gln Ala Gln Gly Arg Asp
 545 550 555 560
 Trp Leu Leu Gly Leu Asn Asn Leu Asp Gly Leu His Arg Ala Leu Gly
 565 570 575
 Leu Ala His Gly Arg Leu Ala Asp Pro Ser Thr Pro Pro Ile Arg Leu
 580 585 590
 Ala Pro Leu Arg Asn Leu Gly Leu Arg Val Leu Val Val Glu Asp Asn
 595 600 605
 Ala Ile Asn Gln Leu Ile Leu Arg Asp Gln Met Glu Ala Leu Gly Cys
 610 615 620
 Ser Val Glu Leu Leu Phe Asp Gly Arg Glu Ala Leu Leu His Cys Gln
 625 630 635 640
 Thr Ala Cys Phe Asp Val Val Leu Thr Asp Ile Asn Met Pro Asn Met
 645 650 655
 Asn Gly Tyr Glu Leu Thr Ala Glu Leu Arg Arg Gln Gly Phe Arg Gln
 660 665 670
 Pro Ile Ile Gly Ala Thr Ala Asn Ala Met Arg Glu Glu Arg Glu Arg
 675 680 685
 Cys Met Ser Ala Gly Met Asn Asp Cys Leu Val Lys Pro Val Asp Leu
 690 695 700
 Asn Ala Leu Gln Asn Cys Leu Ile Asn Ile Leu Lys Val Asp Arg
 705 710 715

65/2

Figure 6T

ORF1-9

SEQ ID NO:27

```

Met Leu Leu Asp Gln His Gly Gln Val Val Leu Ala Thr Asp Ala Glu
 1      5      10      15
Ala Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp Gly Phe
 20      25      30
Gly Phe Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu Phe Gln
 35      40      45
His Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly Ile Gly
 50      55      60
Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser Ala Leu
 65      70      75      80
Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser Ile Glu
 85      90      95
Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu Lys Glu
100      105      110
Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val Ala Leu
115      120      125
Cys Val Leu Arg Arg Ala Asp Ala Val Val Leu Glu Asn Pro Gln
130      135      140
Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp Ala Pro
145      150      155      160
Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser Gly Glu
165      170      175
Glu Leu Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr Thr Pro
180      185      190
Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser Glu Ile
195      200      205
Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu
210      215      220
Ala Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr Met Ser
225      230      235      240
His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu
245      250      255
Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala
260      265      270
Ile Gln His Ser Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp Val Leu
275      280      285
Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys Val Glu
290      295      300
Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe Thr Gly
305      310      315      320
Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser Ala Glu
325      330      335
Leu Pro Leu Arg Met Arg Gly Ala Ala Ser Ile Arg Gln Ile Leu
340      345      350
Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly Tyr Val
355      360      365
Asn Val His Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val Met Leu
370      375      380
Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu Asp Gln
385      390      395      400
Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu His Pro
405      410      415
Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln
420      425      430
Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu Gly Ser
435      440      445
Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln Ala Glu
450      455      460

```

6T/1

Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro Val Arg
 465 470 475 480
 Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg
 485 490 495
 Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser
 500 505 510
 Leu Leu Val Glu Val Leu Leu Leu Glu Gly Ala Pro Met Phe Glu Ala
 515 520 525
 Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met Glu Pro
 530 535 540
 Gln Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu Asn Asn Leu Asp Gly
 545 550 555 560
 Leu His Arg Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp Pro Ser
 565 570 575
 Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu Arg Val
 580 585 590
 Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg Asp Gln
 595 600 605
 Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly Arg Glu
 610 615 620
 Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp Val Val Leu Thr Asp
 625 630 635 640
 Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu Leu Arg
 645 650 655
 Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn Ala Met
 660 665 670
 Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp Cys Leu
 675 680 685
 Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile Asn Ile
 690 695 700
 Leu Lys Val Asp Arg
 705

6T/2

Figure 6U

ORF1-10

SEQ ID NO:28

Met Val Leu Phe Gln His Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr
 1 5 10 15
 His Ile Gly Ile Gly Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu
 20 25 30
 Leu Ala Ser Ala Leu Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu
 35 40 45
 Val Arg Ser Ile Glu Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu
 50 55 60
 Glu Ala Leu Lys Glu Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala
 65 70 75 80
 Ala Pro Val Ala Leu Cys Val Leu Arg Arg Ala Asp Ala Val Val
 85 90 95
 Leu Glu Asn Pro Gln Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile
 100 105 110
 Ala His Asp Ala Pro Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val
 115 120 125
 Lys Cys Ser Gly Glu Glu Leu Glu Thr Glu Ala Gly Leu His Leu His
 130 135 140
 Leu Asn Tyr Thr Pro Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys
 145 150 155 160
 Ala Phe Ser Glu Ile Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala
 165 170 175
 Arg Ala Lys Ser Leu Ala Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe
 180 185 190
 Leu Ala Thr Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu
 195 200 205
 Gly Thr Leu Glu Leu Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala
 210 215 220
 Gly Tyr Leu Lys Ala Ile Gln His Ser Ser Ser Thr Leu Leu Gln Leu
 225 230 235 240
 Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp
 245 250 255
 Leu Glu Cys Val Glu Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val
 260 265 270
 Gln Ser Phe Thr Gly Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr
 275 280 285
 Cys Leu Ser Ala Glu Leu Pro Leu Arg Met Arg Gly Ala Ala Ala Ser
 290 295 300
 Ile Arg Gln Ile Leu Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr
 305 310 315 320
 Asp Asn Gly Tyr Val Asn Val His Leu Lys Ala Ser Val Val Asp Ala
 325 330 335
 Glu Cys Val Met Leu Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile
 340 345 350
 Asn Val Glu Asp Gln Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg
 355 360 365
 Arg Ser Glu His Pro Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser
 370 375 380
 Gln Arg Leu Ala Gln Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu
 385 390 395 400
 Leu Gly Leu Gly Ser Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile
 405 410 415
 Ala Met Gln Ala Glu Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val
 420 425 430
 Leu Ala Pro Val Arg Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser
 435 440 445
 Arg Trp Gly Gly Arg Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu
 450 455 460

64/1

Ala Asp Ala Thr Ser Leu Leu Val Glu Val Leu Leu Leu Glu Gly Ala
 465 470 475 480
 Pro Met Phe Glu Ala Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln
 485 490 495
 Gly Asp Met Glu Pro Gln Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu
 500 505 510
 Asn Asn Leu Asp Gly Leu His Arg Ala Leu Gly Leu Ala His Gly Arg
 515 520 525
 Leu Ala Asp Pro Ser Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn
 530 535 540
 Leu Gly Leu Arg Val Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu
 545 550 555 560
 Ile Leu Arg Asp Gln Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu
 565 570 575
 Phe Asp Gly Arg Glu Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp
 580 585 590
 Val Val Leu Thr Asp Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu
 595 600 605
 Thr Ala Glu Leu Arg Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala
 610 615 620
 Thr Ala Asn Ala Met Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly
 625 630 635 640
 Met Asn Asp Cys Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn
 645 650 655
 Cys Leu Ile Asn Ile Leu Lys Val Asp Arg
 660 665

64/2

Figure 6V

ORF1-11

SEQ ID NO:29

```

Met Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu Ala Asp Ala Ala Asn
1      5      10      15
Glu Ala Lys Thr Leu Phe Leu Ala Thr Met Ser His Glu Ile Arg Thr
20      25      30
Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu Leu Gly Arg Thr Glu
35      40      45
Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala Ile Gln His Ser Ser
50      55      60
Ser Thr Leu Leu Gln Leu Ile Ser Asp Val Leu Asp Val Ser Lys Ile
65      70      75      80
Glu Ala Gly Gln Leu Asp Leu Glu Cys Val Glu Phe Ser Pro Leu Glu
85      90      95
Leu Thr Glu Glu Val Val Gln Ser Phe Thr Gly Ala Ala Gln Ala Lys
100     105     110
Gly Leu Gln Leu Tyr Thr Cys Leu Ser Ala Glu Leu Pro Leu Arg Met
115     120     125
Arg Gly Ala Ala Ala Ser Ile Arg Gln Ile Leu Asn Asn Leu Leu Ser
130     135     140
Asn Ala Val Lys Phe Thr Asp Asn Gly Tyr Val Asn Val His Leu Lys
145     150     155     160
Ala Ser Val Val Asp Ala Glu Cys Val Met Leu Thr Trp Gln Val Asn
165     170     175
Asp Thr Gly Met Gly Ile Asn Val Glu Asp Gln Pro Arg Leu Phe Glu
180     185     190
Pro Phe Tyr Gln Ile Arg Arg Ser Glu His Pro Val Ala Gly Thr Gly
195     200     205
Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln Leu Met Asn Gly Ser
210     215     220
Leu Lys Leu Val Ser Glu Leu Gly Leu Gly Ser Ser Phe Ser Leu Arg
225     230     235     240
Leu Pro Leu Glu Arg Ile Ala Met Gln Ala Glu Pro Gln Asp Leu Ala
245     250     255
Gly Cys Ala Val Gln Val Leu Ala Pro Val Arg Asp Leu Thr Glu Cys
260     265     270
Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg Ala Met Val Ala Thr
275     280     285
Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser Leu Leu Val Glu Val
290     295     300
Leu Leu Leu Glu Gly Ala Pro Met Phe Glu Ala Trp Pro Gly Cys Arg
305     310     315     320
Val Glu Leu Ser Pro Gln Gly Asp Met Glu Pro Gln Ala Gln Gly Arg
325     330     335
Asp Trp Leu Leu Gly Leu Asn Asn Leu Asp Gly Leu His Arg Ala Leu
340     345     350
Gly Leu Ala His Gly Arg Leu Ala Asp Pro Ser Thr Pro Pro Ile Arg
355     360     365
Leu Ala Pro Leu Arg Asn Leu Gly Leu Arg Val Leu Val Val Glu Asp
370     375     380
Asn Ala Ile Asn Gln Leu Ile Leu Arg Asp Gln Met Glu Ala Leu Gly
385     390     395     400
Cys Ser Val Glu Leu Phe Asp Gly Arg Glu Ala Leu Leu His Cys
405     410     415
Gln Thr Ala Cys Phe Asp Val Val Leu Thr Asp Ile Asn Met Pro Asn
420     425     430
Met Asn Gly Tyr Glu Leu Thr Ala Glu Leu Arg Arg Gln Gly Phe Arg
435     440     445
Gln Pro Ile Ile Gly Ala Thr Ala Asn Ala Met Arg Glu Glu Arg Glu
450     455     460

```

6V/1

WO 03/004689

PCT/US02/21431

Arg	Cys	Met	Ser	Ala	Gly	Met	Asn	Asp	Cys	Leu	Val	Lys	Pro	Val	Asp
465					470					475					480
Leu	Asn	Ala	Leu	Gln	Asn	Cys	Leu	Ile	Asn	Ile	Leu	Lys	Val	Asp	Arg
				485					490					495	

6V/2

Figure 7a

ORF3-2

SEQ ID NO:30

atgggatgtta	tacgggagca	tgaggtat	tttgggcgca	tcgctcgaaa	aagcgacaag	60
accacccaga	agtacgacta	tgacgtggtg	cctttgcagc	ggcacttgtt	ggcaaaggaa	120
aacggattag	cggctctatga	gggacgggag	ttttcctttg	ctatgccatt	tctactggct	180
accaagcacg	cgttgagcgc	cgattcctcg	ggagatccgt	tttcgctcgg	tgtattgctc	240
gccaatttct	acggaaagctt	ctggagtgtt	tccgcctatc	ccgcgccaca	gttactgac	300
tttgatcttt	ccggcagcac	ccgcctggca	gtgccgtcga	ttccctccac	agcgcagcgt	360
gacaggttga	gcggaagcta	tccgatgata	gtcgagcgca	ttctggcgcg	cttgccgacc	420
cggccggtgg	gggaggacgc	tcagcgtgtc	cattggatac	gcgctgacg	ctatcgcgac	480
tcggcgctgg	agatgttggg	agtcgcccgg	gttgatctgc	cggaacact	ctgggtggcac	540
gacgagccga	accatctgat	catcgctgcg	agcctgcttg	atctcaggcg	aatcaatgac	600
ttcgaacagt	tggttgagcg	cccggcattc	gattcgtaca	gcctgggtatc	gccggatggc	660
gaggtattgc	tcggcgcggc	ccctgcgacc	ggcctgaggg	atggcctgaa	cctcaccoga	720
cagggggctcg	ccgttcaact	gcgcagccag	cctgagaacg	gctggctcgc	ggtctaccga	780
accgactacg	gcaatttctt	tcgccactcc	cgggtggctgg	tggcaggtct	gctgctgacc	840
ccggcgctgc	tcctggcccg	ttggctcggg	atgcgttggg	acaccagcag	cgctcgtcaac	900
ccggtgcatc	gggcgccaccg	gcaactggtg	gagagcgaca	ccttcagccg	gacgctgata	960
cagaccgcgc	cgggtggctct	ggtggtgctg	accaggatg	accagcaact	ggtgacctgc	1020
aaccacttgg	ccgccagtg	gctgggcggg	cccacggaga	tccttgggct	gacttccaac	1080
tggaaagcttt	tcgatgcgcg	tgggcaggta	ccaggagaca	tctgtatcca	ggtcgggtggg	1140
cgctatttgc	agaccgcctt	cgcgcgacc	cgctatgccg	gcaccgaggg	ggtactgtgc	1200
gtattcaacg	acatcacggg	ccactgcgag	gcggagaccg	cgctgtccaa	tgcgaaagcga	1260
gcagcggatg	ccgccagcca	ggccaagacc	ctgttcctgg	cccgcctgag	ccatgaaatc	1320
cgtactcccc	tgtacggtgt	ccttggcacc	ctggagtgtc	tcgacctgac	cacctgaac	1380
gagcggcaac	gcgcctacct	acgcaccatc	cagagtctgt	ctgcgacgct	catgcaactg	1440
attagcgatg	tgctggatgt	ctcgaagatc	gaagcggggc	agatggctct	gaccttggtc	1500
gccttcaatc	cgtcggacct	agtgcgggaa	gtgcttggca	actttgccgc	cagcgccatg	1560
gccaaggacc	tgcaggtaga	cccgcctgat	actcttgccg	ttgaggcgca	ggtcgcgcac	1620
ggcttcgaag	aaagcgttct	gttcgaggtt	gctggtggct	cggtcggcca	tttcgaagag	1680
ggtgtcgtcg	gcgttgtcga	acaacgcctg	caacgcctgt	ttcagctgca	gcgccgcctt	1740
gtcgcgcacc	tgcacgagga	tgaccggcag	gcgccccgct	ccggcggttcg	gcgacggctc	1800
ggaagcgacc	ctggtcaggt	gcaccacatt	ggcatcggtc	tgcacgcggg	ctctcctgcc	1860
accctcgcgg	ccgcgcctgg	aatggcaaaa	atcgggcaca	gaggatcgat	tggcgtcgtc	1920
cgtaacgtca	atttccaggg	gtcaaaaaaca	agtatctaca	ttcattatag	agatactttc	1980
aaatctagat	ag					1992

Figure 7B

ORF3-3

SEQ ID NO:31

```

atgccatttc tactggctac caagcacgcg ttgagcgccg attcctcggg agatccgttt 60
tcgctcgggtg tattgctcgc caattttctac ggaagcttct ggagtgtttc cgcctatccc 120
gcgccacagt tactgatctt tgatctttcc ggcagcaccg gcctggcagt gccgtcgatt 180
ccctccacag cgcagcggtga cagggtgagc ggaagctatc cgatgatagt cgagcgcatt 240
ctggcgcgct tgcgcacccg gccggtgggg gaggacgctc agcgtgtcca ttggatacgc 300
gtgatcgct atcgcgactc ggcgctggag atgttgggag tcgcccgggt tgatctgccg 360
gaaacactct ggtggcacga cgagccgaac catctgatca tcgctgcgag cctgcttgat 420
ctcaggcgaa tcaatgactt cgaacagttg gttgagcgcc cggcattcga ttcgtacagc 480
ctggtatcgc cggatggcga ggtattgctc ggcgcggccc ctgcgaccgg cctgagggat 540
ggcctgaacc tcacccgaca gggggtcgcc gttcaactgc gcagccagcc tgagaacggc 600
tggtcgcggg tctaccgaac cgactacggc aatttcttcc gccactcccg gtggctgggtg 660
gcagggtcgc tgctgacccc ggcgctgctc ctggccgggt ggctcgggat gcgttggtac 720
accagcagcg tcgtcaaccc ggtgcatcgg gcgcaccggc aactggtgga gagcgacacc 780
ttcagccgga cgctgataca gaccgcgcgg gtggtctctg ttggtgctgac ccaggatgac 840
cagcaactgg tgacctgcaa ccacttgccc gccacgtggc tgggcggggc cacggagatc 900
cttgggctga cttccaactg gaagcttttc gatgcgcgtg ggcagggtacc aggagacatc 960
tgtatccagg tcggtgggcg ctatttgtag accgccttcg cggcgacccg ctatgccggc 1020
accgagggcg tactgtgcgt attcaacgac atcacgggtc actgcgaggc ggagaccgcg 1080
ctgtccaatg cgaagcgagc agcggatgcc gccagccagg ccaagaccct gttcctggcc 1140
cgcatgagcc atgaaatccg tactcccctg tacgggtgtcc ttggcaccct ggagttgctc 1200
gacctgacca ccctgaacga gcggcaacgc gcctacctac gcaccatcca gagttcgtct 1260
gcgacgctca tgcaactgat tagcgatgtg ctggatgtct cgaagatcga agcggggcag 1320
atggctctga ccctggccgc cttcaatccg ctggacctag tgcgggaagt gcttggcaac 1380
tttgccgcca gcgccatggc caaggacctg caggtagacc cgctcgatac tcttgcgctt 1440
gaggcgagg tcgcgcatgg cttcgaagaa agcgttctgt tcgaggttgc tgggtggctc 1500
gtcggccatt tcgaagaggg tgtcgtcggc gttgtcgaac aacgcctgca acgcctgttt 1560
cagctgcagc gccgccttgt cgcgcacctg cacgaggatg accggcaggc gcccgcctcc 1620
ggcgctcggc gacggctcgg aagegacctt ggtcagggtg accacattgg catcgttctg 1680
catcgggact ctctgccac cctcgcggcc gcgcatggaa tggcaaaaat cgggcacaga 1740
ggatcgattg gcgtcgtccg taacgtcaat ttccaggcgt caaaaacaag tatctacatt 1800
cattatagag atactttcaa atctagatag                                     1830

```

Figure 7C

ORF3-4

SEQ ID NO:32

```

atgatatgtcg agcgcattct ggcgcgcttg cgcacccggc cggtaggggga ggacgctcag 60
cgtgtccatt ggatacgcgc tgatcgctat cgcgactcgg cgctggagat gttgggagtc 120
gcccgggttg atctgccgga aacactcttg tggcacgacg agccgaacca tctgatcatc 180
gctgcgagcc tgcttgatct caggcgaatc aatgacttcg aacagttggt tgagcgcccg 240
gcattcgatt cgtacagcct ggtatcgccg gatggcgagg tattgctcgg cgcggccct 300
gcgaccggcc tgagggatgg cctgaacctc acccgacagg gggtcgccgt tcaactgcgc 360
agccagcctg agaacggctg gctcgcggtc taccgaaccg actacggcaa tttctttcgc 420
cactcccggg ggctgggtggc aggtctgctg ctgaccccg cgtgctcct ggccgggttg 480
ctcgggatgc gttggtacac cagcagcgtc gtcaaccgg tgcatcgggc gcaccggcaa 540
ctgggtggaga gcgacacctt cagccggacg ctgatacaga ccgcgccggg ggctctggtg 600
gtgctgaccc aggatgacca gcaactggtg acctgcaacc acttggccgc ccagtggctg 660
ggcgggcccc cggagatcct tgggctgact tccaactgga agcttttcga tgcgcgtggg 720
caggtaccag gagacatctg tatccaggtc ggtgggcgct atttgagac cgccttcgcg 780
gcgaccgcgt atgccggcac cgaggcggtg ctgtgcgtat tcaacgacat caccgtccac 840
tgcgaggcgg agaccgcgct gtccaatgcg aagcgagcag cggatgccgc cagccaggcc 900
aagaccctgt tcctggcccc .catgagccat gaaatccgta ctcccctgta cgggtgcctt 960
ggcaccctgg agttgctcga cctgaccacc ctgaacgagc ggcaacgcgc ctacctacgc 1020
accatccaga gttcgtctgc gacgctcatg caactgatta gcgatgtgct ggatgtctcg 1080
aagatcgaag cggggcagat ggctctgacc ctggccgcct tcaatccgct ggacctagtg 1140
cgggaagtgc ttggcaactt tgccgccagc gccatggcca aggacctgca ggtagacccg 1200
ctcgatactc ttgcgcttga ggcgaggctc gcgcatggct tcgaagaaag cgttctgttc 1260
gaggttgctg gtggctcggg cgccatttc gaagagggtg tcgtcggcgt tgtcgaacaa 1320
cgctgcaac gcctgtttca gctgcagcgc cgccttgctg cgcacctgca cgaggatgac 1380
cggcaggcgc cccgctccgg cgttcggcga cggctcggaa gcgaccctgg tcagggtgcac 1440
cacattggca tcgttctgca tcgggactct cctgccacc tcgcggccgc gcattggaatg 1500
gcaaaaaatcg ggcacagagg atcgattggc gtcgctccgta acgtcaattt ccaggcgtca 1560
aaaacaagta tctacattca ttatagagat actttcaaat ctagatag 1608

```

Figure 7D

ORF3-5

SEQ ID NO:33

```

atgttgggag tcgcccgggt tgatctgccg gaaacactct ggtggcacga cgagccgaac 60
catctgatca tcgctgcgag cctgcttgat ctcaggcgaa tcaatgactt cgaacagttg 120
gttgagcgcc cggcattcga ttctgtacagc ctgggtatcg cggatggcga ggtattgctc 180
ggcgcgggcc ctgcgaccgg cctgagggat ggcctgaacc tcaccgcaca gggggtcgcc 240
gttcaactgc gcagccagcc tgagaacggc tggctcgccg tctaccgaac cgactacggc 300
aatttcttcc gccactcccg gtggctgggt gcaggtctgc tgetgacccc ggcgctgctc 360
ctggccgggt ggctcgggat gcgttggtac accagcagcg tctcaacccc ggtgcatcgg 420
gcgcaccggc aactgggtga gagcgacacc ttcagccgga cgctgataca gaccgcgcgc 480
gtggctctgg tgggtgctgac ccaggatgac cagcaactgg tgacctgcaa ccacttgccc 540
gcccagtggt tgggcgggcc cacggagatc cttgggctga cttccaactg gaagcttttc 600
gatgcgcgtg ggcaggtacc aggagacatc tgtatccagg tcggtgggcg ctatttgacg 660
accgccttcg cggcgacccg ctatgccggc accgagggcg tactgtgcgt attcaacgac 720
atcacggtcc actgcgaggg ggagaccgcg ctgtccaatg cgaagcgagc agcggatgcc 780
gccagccagg ccaagaccct gttcctggcc cgcagtgacc atgaaatccg tactcccctg 840
tacggtgtcc ttggcacccct ggagttgctc gacctgacca ccctgaacga gcggcaacgc 900
gcctacctac gcaccatcca gagttcgtct gcgacgctca tgcaactgat tagcgatgtg 960
ctggatgtct cgaagatcga agcggggcag atggctctga ccctggccgc cttcaatccg 1020
ctggacctag tgcgggaagt gcttggaac tttgccgcca gcgccatggc caaggacctg 1080
caggtagacc cgctcgatac tcttgcgctt gaggcgagg tcgcgcatgg cttcgaagaa 1140
agcgttctgt tcgaggttgc tgggtggctc gtcggccatt tcgaagaggg tgtcgtcgcc 1200
gttgtcgaac aacgcctgca acgcctgtt cagctgcagc gccgccttgt cgcgcacctg 1260
cacgaggatg accggcaggg gccccgctcc ggcgttcggc gacggctcgg aagcgaccct 1320
ggtcaggtgc accacattgg catcgttctg catcgggact ctcctgccac ctcgcgggcc 1380
gcgcattgaa tggcaaaaat cgggcacaga ggatcgattg gcgtcgtccg taacgtcaat 1440
ttccaggcgt caaaaacaag tatctacatt cattatagag atactttcaa atctagatag 1500

```

Figure 7B

ORF3-6

SEQ ID NO:34

atgcgttggg	acaccagcag	cgctcgtcaac	ccggtgcatc	gggcgcaccg	gcaactgggtg	60
gagagcgaca	ccttcagccg	gacgctgata	cagaccgcgc	cggtggctct	ggtgggtgctg	120
accaggatg	accagcaact	ggtgacctgc	aaccacttgg	ccgccagtg	gctgggcggg	180
cccacggaga	tccttgggct	gacttccaac	tgggaagctt	tcgatgcgcg	tgggcaggta	240
ccaggagaca	tctgtatcca	ggtcgggtggg	cgctatttgc	agaccgcctt	cgcggcgacc	300
cgctatgccg	gcaccgaggc	ggtactgtgc	gtattcaacg	acatcacggg	ccactgcgag	360
gcggagaccg	cgctgtccaa	tgcaagcgca	gcagcggatg	ccgccagcca	ggccaagacc	420
ctgttcctgg	cccgcctgag	ccatgaaatc	cgtaactccc	tgtacgggtg	ccttggcacc	480
ctggagttgc	tcgacctgac	cacctgaac	gagcggcaac	gcgcctacct	acgcaccatc	540
cagagttcgt	ctgcgacgct	catgcaactg	attagcgatg	tgctggatgt	ctcgaagatc	600
gaagcggggc	agatggctct	gacctggcc	gccttcaatc	cgctggacct	agtgcgggaa	660
gtgcttggca	actttgccgc	cagcgccatg	gccaaaggacc	tgcaaggtaga	cccgctcgat	720
actcttgccg	ttgaggcgca	ggtcgcgcgc	ggcttcgaag	aaagcggtct	gttcgaggtt	780
gctgggtggc	cggtcggcca	tttcgaagag	ggtgtcgtcg	gcgttgctga	acaacgcctg	840
caacgcctgt	ttcagctgca	gcgccgcctt	gtcgcgcacc	tgcaaggagga	tgaccggcag	900
gcgccccgct	ccggcggttcg	gcgacggctc	ggaagcgacc	ctggtcagggt	gcaccacatt	960
ggcatcgttc	tgcatcgggg	ctctcctgcc	accctcgccg	ccgcgcctgg	aatggcaaaa	1020
atcgggcaca	gaggatcgat	tggcgctcgtc	cgtaacgtca	atttcagggc	gtcaaaaaca	1080
agtatctaca	ttcattatag	agatactttc	aaatctagat	ag		1122

Figure 7F

ORF3-2

SEQ ID NO:35

Met Asp Val Ile Arg Glu His Glu Val Phe Leu Gly Arg Ile Ala Arg
 1 5 10 15
 Lys Ser Asp Lys Thr Thr Gln Lys Tyr Asp Tyr Asp Val Val Pro Leu
 20 25 30
 Gln Arg His Leu Leu Ala Lys Glu Asn Gly Leu Ala Val Tyr Glu Gly
 35 40 45
 Arg Glu Phe Ser Phe Ala Met Pro Phe Leu Leu Ala Thr Lys His Ala
 50 55 60
 Leu Ser Ala Asp Ser Ser Gly Asp Pro Phe Ser Leu Gly Val Leu Leu
 65 70 75 80
 Ala Asn Phe Tyr Gly Ser Phe Trp Ser Val Ser Ala Tyr Pro Ala Pro
 85 90 95
 Gln Leu Leu Ile Phe Asp Leu Ser Gly Ser Thr Arg Leu Ala Val Pro
 100 105 110
 Ser Ile Pro Ser Thr Ala Gln Arg Asp Arg Leu Ser Gly Ser Tyr Pro
 115 120 125
 Met Ile Val Glu Arg Ile Leu Ala Arg Leu Arg Thr Arg Pro Val Gly
 130 135 140
 Glu Asp Ala Gln Arg Val His Trp Ile Arg Ala Asp Arg Tyr Arg Asp
 145 150 155 160
 Ser Ala Leu Glu Met Leu Gly Val Ala Arg Val Asp Leu Pro Glu Thr
 165 170 175
 Leu Trp Trp His Asp Glu Pro Asn His Leu Ile Ile Ala Ala Ser Leu
 180 185 190
 Leu Asp Leu Arg Arg Ile Asn Asp Phe Glu Gln Leu Val Glu Arg Pro
 195 200 205
 Ala Phe Asp Ser Tyr Ser Leu Val Ser Pro Asp Gly Glu Val Leu Leu
 210 215 220
 Gly Ala Ala Pro Ala Thr Gly Leu Arg Asp Gly Leu Asn Leu Thr Arg
 225 230 235 240
 Gln Gly Val Ala Val Gln Leu Arg Ser Gln Pro Glu Asn Gly Trp Leu
 245 250 255
 Ala Val Tyr Arg Thr Asp Tyr Gly Asn Phe Phe Arg His Ser Arg Trp
 260 265 270
 Leu Val Ala Gly Leu Leu Leu Thr Pro Ala Leu Leu Leu Ala Gly Trp
 275 280 285
 Leu Gly Met Arg Trp Tyr Thr Ser Ser Val Val Asn Pro Val His Arg
 290 295 300
 Ala His Arg Gln Leu Val Glu Ser Asp Thr Phe Ser Arg Thr Leu Ile
 305 310 315 320
 Gln Thr Ala Pro Val Ala Leu Val Val Leu Thr Gln Asp Asp Gln Gln
 325 330 335
 Leu Val Thr Cys Asn His Leu Ala Ala Gln Trp Leu Gly Gly Pro Thr
 340 345 350
 Glu Ile Leu Gly Leu Thr Ser Asn Trp Lys Leu Phe Asp Ala Arg Gly
 355 360 365
 Gln Val Pro Gly Asp Ile Cys Ile Gln Val Gly Gly Arg Tyr Leu Gln
 370 375 380
 Thr Ala Phe Ala Ala Thr Arg Tyr Ala Gly Thr Glu Ala Val Leu Cys
 385 390 395 400
 Val Phe Asn Asp Ile Thr Val His Cys Glu Ala Glu Thr Ala Leu Ser
 405 410 415
 Asn Ala Lys Arg Ala Ala Asp Ala Ala Ser Gln Ala Lys Thr Leu Phe
 420 425 430
 Leu Ala Arg Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Val Leu
 435 440 445
 Gly Thr Leu Glu Leu Leu Asp Leu Thr Thr Leu Asn Glu Arg Gln Arg
 450 455 460

7F/1

Ala	Tyr	Leu	Arg	Thr	Ile	Gln	Ser	Ser	Ser	Ala	Thr	Leu	Met	Gln	Leu
465					470					475					480
Ile	Ser	Asp	Val	Leu	Asp	Val	Ser	Lys	Ile	Glu	Ala	Gly	Gln	Met	Ala
				485					490					495	
Leu	Thr	Leu	Ala	Ala	Phe	Asn	Pro	Leu	Asp	Leu	Val	Arg	Glu	Val	Leu
			500					505					510		
Gly	Asn	Phe	Ala	Ala	Ser	Ala	Met	Ala	Lys	Asp	Leu	Gln	Val	Asp	Pro
		515					520					525			
Leu	Asp	Thr	Leu	Ala	Leu	Glu	Ala	Gln	Val	Ala	His	Gly	Phe	Glu	Glu
	530					535					540				
Ser	Val	Leu	Phe	Glu	Val	Ala	Gly	Gly	Ser	Val	Gly	His	Phe	Glu	Glu
545					550					555					560
Gly	Val	Val	Gly	Val	Val	Glu	Gln	Arg	Leu	Gln	Arg	Leu	Phe	Gln	Leu
			565						570					575	
Gln	Arg	Arg	Leu	Val	Ala	His	Leu	His	Glu	Asp	Asp	Arg	Gln	Ala	Pro
			580					585					590		
Arg	Ser	Gly	Val	Arg	Arg	Arg	Leu	Gly	Ser	Asp	Pro	Gly	Gln	Val	His
		595					600					605			
His	Ile	Gly	Ile	Val	Leu	His	Arg	Asp	Ser	Pro	Ala	Thr	Leu	Ala	Ala
	610					615					620				
Ala	His	Gly	Met	Ala	Lys	Ile	Gly	His	Arg	Gly	Ser	Ile	Gly	Val	Val
625					630					635					640
Arg	Asn	Val	Asn	Phe	Gln	Ala	Ser	Lys	Thr	Ser	Ile	Tyr	Ile	His	Tyr
			645						650					655	
Arg	Asp	Thr	Phe	Lys	Ser	Arg									
			660												

7F/2

Figure 7G

ORF 3-3

SEQ ID NO:36

```

Met Pro Phe Leu Leu Ala Thr Lys His Ala Leu Ser Ala Asp Ser Ser
1      5      10      15
Gly Asp Pro Phe Ser Leu Gly Val Leu Leu Ala Asn Phe Tyr Gly Ser
20      25      30
Phe Trp Ser Val Ser Ala Tyr Pro Ala Pro Gln Leu Leu Ile Phe Asp
35      40      45
Leu Ser Gly Ser Thr Arg Leu Ala Val Pro Ser Ile Pro Ser Thr Ala
50      55      60
Gln Arg Asp Arg Leu Ser Gly Ser Tyr Pro Met Ile Val Glu Arg Ile
65      70      75      80
Leu Ala Arg Leu Arg Thr Arg Pro Val Gly Glu Asp Ala Gln Arg Val
85      90      95
His Trp Ile Arg Ala Asp Arg Tyr Arg Asp Ser Ala Leu Glu Met Leu
100     105     110
Gly Val Ala Arg Val Asp Leu Pro Glu Thr Leu Trp Trp His Asp Glu
115     120     125
Pro Asn His Leu Ile Ile Ala Ser Leu Leu Asp Leu Arg Arg Ile
130     135     140
Asn Asp Phe Glu Gln Leu Val Glu Arg Pro Ala Phe Asp Ser Tyr Ser
145     150     155     160
Leu Val Ser Pro Asp Gly Glu Val Leu Leu Gly Ala Ala Pro Ala Thr
165     170     175
Gly Leu Arg Asp Gly Leu Asn Leu Thr Arg Gln Gly Val Ala Val Gln
180     185     190
Leu Arg Ser Gln Pro Glu Asn Gly Trp Leu Ala Val Tyr Arg Thr Asp
195     200     205
Tyr Gly Asn Phe Phe Arg His Ser Arg Trp Leu Val Ala Gly Leu Leu
210     215     220
Leu Thr Pro Ala Leu Leu Leu Ala Gly Trp Leu Gly Met Arg Trp Tyr
225     230     235     240
Thr Ser Ser Val Val Asn Pro Val His Arg Ala His Arg Gln Leu Val
245     250     255
Glu Ser Asp Thr Phe Ser Arg Thr Leu Ile Gln Thr Ala Pro Val Ala
260     265     270
Leu Val Val Leu Thr Gln Asp Asp Gln Gln Leu Val Thr Cys Asn His
275     280     285
Leu Ala Ala Gln Trp Leu Gly Gly Pro Thr Glu Ile Leu Gly Leu Thr
290     295     300
Ser Asn Trp Lys Leu Phe Asp Ala Arg Gly Gln Val Pro Gly Asp Ile
305     310     315     320
Cys Ile Gln Val Gly Gly Arg Tyr Leu Gln Thr Ala Phe Ala Ala Thr
325     330     335
Arg Tyr Ala Gly Thr Glu Ala Val Leu Cys Val Phe Asn Asp Ile Thr
340     345     350
Val His Cys Glu Ala Glu Thr Ala Leu Ser Asn Ala Lys Arg Ala Ala
355     360     365
Asp Ala Ala Ser Gln Ala Lys Thr Leu Phe Leu Ala Arg Met Ser His
370     375     380
Glu Ile Arg Thr Pro Leu Tyr Gly Val Leu Gly Thr Leu Glu Leu Leu
385     390     395     400
Asp Leu Thr Thr Leu Asn Glu Arg Gln Arg Ala Tyr Leu Arg Thr Ile
405     410     415
Gln Ser Ser Ser Ala Thr Leu Met Gln Leu Ile Ser Asp Val Leu Asp
420     425     430
Val Ser Lys Ile Glu Ala Gly Gln Met Ala Leu Thr Leu Ala Ala Phe
435     440     445
Asn Pro Leu Asp Leu Val Arg Glu Val Leu Gly Asn Phe Ala Ala Ser
450     455     460

```

7G/1

Ala Met Ala Lys Asp Leu Gln Val Asp Pro Leu Asp Thr Leu Ala Leu
 465 470 475 480
 Glu Ala Gln Val Ala His Gly Phe Glu Glu Ser Val Leu Phe Glu Val
 485 490 495
 Ala Gly Gly Ser Val Gly His Phe Glu Glu Gly Val Val Gly Val Val
 500 505 510
 Glu Gln Arg Leu Gln Arg Leu Phe Gln Leu Gln Arg Arg Leu Val Ala
 515 520 525
 His Leu His Glu Asp Asp Arg Gln Ala Pro Arg Ser Gly Val Arg Arg
 530 535 540
 Arg Leu Gly Ser Asp Pro Gly Gln Val His His Ile Gly Ile Val Leu
 545 550 555 560
 His Arg Asp Ser Pro Ala Thr Leu Ala Ala Ala His Gly Met Ala Lys
 565 570 575
 Ile Gly His Arg Gly Ser Ile Gly Val Val Arg Asn Val Asn Phe Gln
 580 585 590
 Ala Ser Lys Thr Ser Ile Tyr Ile His Tyr Arg Asp Thr Phe Lys Ser
 595 600 605
 Arg

7/6/2

Figure 7H

ORF3-4

SEQ ID NO:37

Met Ile Val Glu Arg Ile Leu Ala Arg Leu Arg Thr Arg Pro Val Gly
 1 5 10 15
 Glu Asp Ala Gln Arg Val His Trp Ile Arg Ala Asp Arg Tyr Arg Asp
 20 25 30
 Ser Ala Leu Glu Met Leu Gly Val Ala Arg Val Asp Leu Pro Glu Thr
 35 40 45
 Leu Trp Trp His Asp Glu Pro Asn His Leu Ile Ile Ala Ala Ser Leu
 50 55 60
 Leu Asp Leu Arg Arg Ile Asn Asp Phe Glu Gln Leu Val Glu Arg Pro
 65 70 75 80
 Ala Phe Asp Ser Tyr Ser Leu Val Ser Pro Asp Gly Glu Val Leu Leu
 85 90 95
 Gly Ala Ala Pro Ala Thr Gly Leu Arg Asp Gly Leu Asn Leu Thr Arg
 100 105 110
 Gln Gly Val Ala Val Gln Leu Arg Ser Gln Pro Glu Asn Gly Trp Leu
 115 120 125
 Ala Val Tyr Arg Thr Asp Tyr Gly Asn Phe Phe Arg His Ser Arg Trp
 130 135 140
 Leu Val Ala Gly Leu Leu Thr Pro Ala Leu Leu Leu Ala Gly Trp
 145 150 155 160
 Leu Gly Met Arg Trp Tyr Thr Ser Ser Val Val Asn Pro Val His Arg
 165 170 175
 Ala His Arg Gln Leu Val Glu Ser Asp Thr Phe Ser Arg Thr Leu Ile
 180 185 190
 Gln Thr Ala Pro Val Ala Leu Val Val Leu Thr Gln Asp Asp Gln Gln
 195 200 205
 Leu Val Thr Cys Asn His Leu Ala Ala Gln Trp Leu Gly Gly Pro Thr
 210 215 220
 Glu Ile Leu Gly Leu Thr Ser Asn Trp Lys Leu Phe Asp Ala Arg Gly
 225 230 235 240
 Gln Val Pro Gly Asp Ile Cys Ile Gln Val Gly Gly Arg Tyr Leu Gln
 245 250 255
 Thr Ala Phe Ala Ala Thr Arg Tyr Ala Gly Thr Glu Ala Val Leu Cys
 260 265 270
 Val Phe Asn Asp Ile Thr Val His Cys Glu Ala Glu Thr Ala Leu Ser
 275 280 285
 Asn Ala Lys Arg Ala Ala Asp Ala Ala Ser Gln Ala Lys Thr Leu Phe
 290 295 300
 Leu Ala Arg Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Val Leu
 305 310 315 320
 Gly Thr Leu Glu Leu Leu Asp Leu Thr Thr Leu Asn Glu Arg Gln Arg
 325 330 335
 Ala Tyr Leu Arg Thr Ile Gln Ser Ser Ser Ala Thr Leu Met Gln Leu
 340 345 350
 Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Met Ala
 355 360 365
 Leu Thr Leu Ala Ala Phe Asn Pro Leu Asp Leu Val Arg Glu Val Leu
 370 375 380
 Gly Asn Phe Ala Ala Ser Ala Met Ala Lys Asp Leu Gln Val Asp Pro
 385 390 395 400
 Leu Asp Thr Leu Ala Leu Glu Ala Gln Val Ala His Gly Phe Glu Glu
 405 410 415
 Ser Val Leu Phe Glu Val Ala Gly Gly Ser Val Gly His Phe Glu Glu
 420 425 430
 Gly Val Val Gly Val Val Glu Gln Arg Leu Gln Arg Leu Phe Gln Leu
 435 440 445
 Gln Arg Arg Leu Val Ala His Leu His Glu Asp Asp Arg Gln Ala Pro
 450 455 460

74/1

Arg Ser Gly Val Arg Arg Arg Leu Gly Ser Asp Pro Gly Gln Val His
465 470 475 480
His Ile Gly Ile Val Leu His Arg Asp Ser Pro Ala Thr Leu Ala Ala
485 490 495
Ala His Gly Met Ala Lys Ile Gly His Arg Gly Ser Ile Gly Val Val
500 505 510
Arg Asn Val Asn Phe Gln Ala Ser Lys Thr Ser Ile Tyr Ile His Tyr
515 520 525
Arg Asp Thr Phe Lys Ser Arg
530 535

74/2

Figure 7I

ORF3-5

SEQ ID NO:38

Met Leu Gly Val Ala Arg Val Asp Leu Pro Glu Thr Leu Trp Trp His
 1 5 10 15
 Asp Glu Pro Asn His Leu Ile Ile Ala Ala Ser Leu Leu Asp Leu Arg
 20 25 30
 Arg Ile Asn Asp Phe Glu Gln Leu Val Glu Arg Pro Ala Phe Asp Ser
 35 40 45
 Tyr Ser Leu Val Ser Pro Asp Gly Glu Val Leu Leu Gly Ala Ala Pro
 50 55 60
 Ala Thr Gly Leu Arg Asp Gly Leu Asn Leu Thr Arg Gln Gly Val Ala
 65 70 75 80
 Val Gln Leu Arg Ser Gln Pro Glu Asn Gly Trp Leu Ala Val Tyr Arg
 85 90 95
 Thr Asp Tyr Gly Asn Phe Phe Arg His Ser Arg Trp Leu Val Ala Gly
 100 105 110
 Leu Leu Leu Thr Pro Ala Leu Leu Leu Ala Gly Trp Leu Gly Met Arg
 115 120 125
 Trp Tyr Thr Ser Ser Val Val Asn Pro Val His Arg Ala His Arg Gln
 130 135 140
 Leu Val Glu Ser Asp Thr Phe Ser Arg Thr Leu Ile Gln Thr Ala Pro
 145 150 155 160
 Val Ala Leu Val Val Leu Thr Gln Asp Asp Gln Gln Leu Val Thr Cys
 165 170 175
 Asn His Leu Ala Ala Gln Trp Leu Gly Gly Pro Thr Glu Ile Leu Gly
 180 185 190
 Leu Thr Ser Asn Trp Lys Leu Phe Asp Ala Arg Gly Gln Val Pro Gly
 195 200 205
 Asp Ile Cys Ile Gln Val Gly Gly Arg Tyr Leu Gln Thr Ala Phe Ala
 210 215 220
 Ala Thr Arg Tyr Ala Gly Thr Glu Ala Val Leu Cys Val Phe Asn Asp
 225 230 235 240
 Ile Thr Val His Cys Glu Ala Glu Thr Ala Leu Ser Asn Ala Lys Arg
 245 250 255
 Ala Ala Asp Ala Ala Ser Gln Ala Lys Thr Leu Phe Leu Ala Arg Met
 260 265 270
 Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Val Leu Gly Thr Leu Glu
 275 280 285
 Leu Leu Asp Leu Thr Thr Leu Asn Glu Arg Gln Arg Ala Tyr Leu Arg
 290 295 300
 Thr Ile Gln Ser Ser Ser Ala Thr Leu Met Gln Leu Ile Ser Asp Val
 305 310 315 320
 Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Met Ala Leu Thr Leu Ala
 325 330 335
 Ala Phe Asn Pro Leu Asp Leu Val Arg Glu Val Leu Gly Asn Phe Ala
 340 345 350
 Ala Ser Ala Met Ala Lys Asp Leu Gln Val Asp Pro Leu Asp Thr Leu
 355 360 365
 Ala Leu Glu Ala Gln Val Ala His Gly Phe Glu Glu Ser Val Leu Phe
 370 375 380
 Glu Val Ala Gly Gly Ser Val Gly His Phe Glu Glu Gly Val Val Gly
 385 390 395 400
 Val Val Glu Gln Arg Leu Gln Arg Leu Phe Gln Leu Gln Arg Arg Leu
 405 410 415
 Val Ala His Leu His Glu Asp Asp Arg Gln Ala Pro Arg Ser Gly Val
 420 425 430
 Arg Arg Arg Leu Gly Ser Asp Pro Gly Gln Val His His Ile Gly Ile
 435 440 445
 Val Leu His Arg Asp Ser Pro Ala Thr Leu Ala Ala Ala His Gly Met
 450 455 460

7I/1

Ala Lys Ile Gly His Arg Gly Ser Ile Gly Val Val Arg Asn Val Asn
465 470 475 480
Phe Gln Ala Ser Lys Thr Ser Ile Tyr Ile His Tyr Arg Asp Thr Phe
485 490 495
Lys Ser Arg

71/2

Figure. 7J

ORF3-6

SEQ ID NO:39

Met Arg Trp Tyr Thr Ser Ser Val Val Asn Pro Val His Arg Ala His
 1 5 10 15
 Arg Gln Leu Val Glu Ser Asp Thr Phe Ser Arg Thr Leu Ile Gln Thr
 20 25 30
 Ala Pro Val Ala Leu Val Val Leu Thr Gln Asp Asp Gln Gln Leu Val
 35 40 45
 Thr Cys Asn His Leu Ala Ala Gln Trp Leu Gly Gly Pro Thr Glu Ile
 50 55 60
 Leu Gly Leu Thr Ser Asn Trp Lys Leu Phe Asp Ala Arg Gly Gln Val
 65 70 75 80
 Pro Gly Asp Ile Cys Ile Gln Val Gly Gly Arg Tyr Leu Gln Thr Ala
 85 90 95
 Phe Ala Ala Thr Arg Tyr Ala Gly Thr Glu Ala Val Leu Cys Val Phe
 100 105 110
 Asn Asp Ile Thr Val His Cys Glu Ala Glu Thr Ala Leu Ser Asn Ala
 115 120 125
 Lys Arg Ala Ala Asp Ala Ala Ser Gln Ala Lys Thr Leu Phe Leu Ala
 130 135 140
 Arg Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Val Leu Gly Thr
 145 150 155 160
 Leu Glu Leu Leu Asp Leu Thr Thr Leu Asn Glu Arg Gln Arg Ala Tyr
 165 170 175
 Leu Arg Thr Ile Gln Ser Ser Ser Ala Thr Leu Met Gln Leu Ile Ser
 180 185 190
 Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Met Ala Leu Thr
 195 200 205
 Leu Ala Ala Phe Asn Pro Leu Asp Leu Val Arg Glu Val Leu Gly Asn
 210 215 220
 Phe Ala Ala Ser Ala Met Ala Lys Asp Leu Gln Val Asp Pro Leu Asp
 225 230 235 240
 Thr Leu Ala Leu Glu Ala Gln Val Ala His Gly Phe Glu Glu Ser Val
 245 250 255
 Leu Phe Glu Val Ala Gly Gly Ser Val Gly His Phe Glu Glu Gly Val
 260 265 270
 Val Gly Val Val Glu Gln Arg Leu Gln Arg Leu Phe Gln Leu Gln Arg
 275 280 285
 Arg Leu Val Ala His Leu His Glu Asp Asp Arg Gln Ala Pro Arg Ser
 290 295 300
 Gly Val Arg Arg Arg Leu Gly Ser Asp Pro Gly Gln Val His His Ile
 305 310 315 320
 Gly Ile Val Leu His Arg Asp Ser Pro Ala Thr Leu Ala Ala Ala His
 325 330 335
 Gly Met Ala Lys Ile Gly His Arg Gly Ser Ile Gly Val Val Arg Asn
 340 345 350
 Val Asn Phe Gln Ala Ser Lys Thr Ser Ile Tyr Ile His Tyr Arg Asp
 355 360 365
 Thr Phe Lys Ser Arg
 370

SEQUENCE LISTING

<110> The General Hospital Corporation

<120> Regulators of Biofilm Formation and Uses
Thereof

<130> 00786/398WO4

<150> US 60/373,233

<151> 2002-04-16

<150> US 60/303,286

<151> 2001-07-06

<160> 39

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1200

<212> DNA

<213> Pseudomonas aeruginosa PA14

<400> 1

```

atgagctgga aatcctatcg ggtgctggtg gtcgaagatc agccgtttca gcgcgaatac 60
ctgctcaacc tgtttcgga gcgcggcgtg cagtacctgg taggtgccgg cgacggcgcg 120
gaggcggttg gctgcctgaa gcaggacagg ttcgacctga tctcagcga tctgatgatg 180
ccgggcatgg atggtatcca aatgatcctg caactgccgt atctcaagca tcgtccgaag 240
ctggcgctga tgagctctc gtcgcagcgg atgatgctca gtgccagccg ggtcgcccag 300
agtctcggct tgtcggtaat cgacctgttg cccaagccga ctctgcccaa ggccatcggc 360
caacttctgg aacacctgga aagatgcctc aggcagaagc tggagccgga aaccgacgag 420
actccgcatg ggcgcacggc gttgctggat gccctgcata acgagcaact ggtgacctgg 480
ttccaggcta agaaatccct ccacaccggg cgcatagtcg gcgccgaggg gttgatacgc 540
tggagccacc cgcagcatgg cctggtgctg ccagctgtt tcatgagtga tgtcgacgct 600
accggtctgc acgaggcggt gctctggcgc gtgctcgaac agaccctgaa cgcccaggaa 660
tcgtggcgca gggcggggta cgagattccg gtttcgggta atctgccgcc gcacctgctc 720
gataaccagg aacttccgga tcgactctat gagtacgtcg gcgctcgagg ggcttgtacc 780
agctcactat gtttcgagtt gaccgagagc agtgtcaca ctctgtcaag taactactat 840
gcaggtgcct gtcgcttgcg catgaaaggg ttcggattgg ccaggacga ctttggccag 900
ggttacagct cgttctataa cctgggtcac acgcctttca cggagctgaa gatcgaccgc 960
tccctagtc agggatgctg agaggataac ggcctcaatg cagctgtcat cagttgtatt 1020
gagttgggtc accgcctgaa tctcgacgtg gtggccgaag gcgtggagac ctgcgaggaa 1080
ctgaatcttc ttcgtcgtct tggctgcgac cggcgcgagg gtttcctgat ttctaaggca 1140
gtgtctgctc gtaggttcca gcggcagtta agggaggacg gccccagcct ccttgtttta 1200

```

<210> 2

<211> 399

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 2

```

Met Ser Trp Lys Ser Tyr Arg Val Leu Val Val Glu Asp Gln Pro Phe
 1             5             10             15
Gln Arg Glu Tyr Leu Leu Asn Leu Phe Arg Glu Arg Gly Val Gln Tyr
          20          25          30

```

Leu Val Gly Ala Gly Asp Gly Ala Glu Ala Leu Arg Cys Leu Lys Gln
 35 40 45
 Asp Arg Phe Asp Leu Ile Leu Ser Asp Leu Met Met Pro Gly Met Asp
 50 55 60
 Gly Ile Gln Met Ile Leu Gln Leu Pro Tyr Leu Lys His Arg Pro Lys
 65 70 75 80
 Leu Ala Leu Met Ser Ser Ser Ser Gln Arg Met Met Leu Ser Ala Ser
 85 90 95
 Arg Val Ala Gln Ser Leu Gly Leu Ser Val Ile Asp Leu Leu Pro Lys
 100 105 110
 Pro Thr Leu Pro Lys Ala Ile Gly Gln Leu Leu Glu His Leu Glu Arg
 115 120 125
 Cys Leu Arg Gln Lys Leu Glu Pro Glu Thr Asp Glu Thr Pro His Gly
 130 135 140
 Arg Thr Ala Leu Leu Asp Ala Leu His Asn Glu Gln Leu Val Thr Trp
 145 150 155 160
 Phe Gln Ala Lys Lys Ser Leu His Thr Gly Arg Ile Val Gly Ala Glu
 165 170 175
 Ala Leu Ile Arg Trp Ser His Pro Gln His Gly Leu Leu Leu Pro Ser
 180 185 190
 Cys Phe Met Ser Asp Val Asp Ala Thr Gly Leu His Glu Ala Leu Leu
 195 200 205
 Trp Arg Val Leu Glu Gln Thr Leu Asn Ala Gln Glu Ser Trp Arg Arg
 210 215 220
 Ala Gly Tyr Glu Ile Pro Val Ser Val Asn Leu Pro Pro His Leu Leu
 225 230 235 240
 Asp Asn Gln Glu Leu Pro Asp Arg Leu Tyr Glu Tyr Val Gly Ala Arg
 245 250 255
 Gly Ala Cys Thr Ser Ser Leu Cys Phe Glu Leu Thr Glu Ser Ser Val
 260 265 270
 Thr Thr Leu Ser Ser Asn Tyr Tyr Ala Gly Ala Cys Arg Leu Arg Met
 275 280 285
 Lys Gly Phe Gly Leu Ala Gln Asp Asp Phe Gly Gln Gly Tyr Ser Ser
 290 295 300
 Phe Tyr Asn Leu Val Thr Thr Pro Phe Thr Glu Leu Lys Ile Asp Arg
 305 310 315 320
 Ser Leu Val Gln Gly Cys Val Glu Asp Asn Gly Leu Asn Ala Ala Val
 325 330 335
 Ile Ser Cys Ile Glu Leu Gly His Arg Leu Asn Leu Asp Val Val Ala
 340 345 350
 Glu Gly Val Glu Thr Cys Glu Glu Leu Asn Leu Leu Arg Arg Leu Gly
 355 360 365
 Cys Asp Arg Ala Gln Gly Phe Leu Ile Ser Lys Ala Val Ser Ala Arg
 370 375 380
 Glu Phe Glu Arg Gln Leu Arg Glu Asp Gly Pro Ser Leu Leu Val
 385 390 395

<210> 3

<211> 1416

<212> DNA

<213> Pseudomonas aeruginosa PA14

<400> 3

atgagccatg aaatccgcac acctctgtac ggcattgcttg gcacgcttga gctgcttggg 60
 cgtaccgagc tgagtcggca gcaggccggt tacctaaagg caatccagca ttctcgtcg 120
 accctgctgc aactgatcag cgatgtgctt gacgtatcca agatagaggc cggccaactg 180
 gacctagagt gcgtggaatt ctccccgctg gaattgaccg aagaggtcgt gcagtcgttc 240


```

accggtgccg cgcaggccaa ggggctgcag ttgtatacct gcctctctgc ggagctgccg 300
ctgcgcgatgc gggggggccgc ggcgtcgatc cggcagattc tcaacaacct gctgagcaac 360
gcggtgaagt tcaccgacaa tggctatgtc aacgtccacc tgaaggccag cgtggtcgat 420
gccgaatgtg tgatgctgac ctggcagggtc aacgataccg gcatggggat caacgtcgag 480
gatcagccgc gtctgttcga accgttctac cagatacgcc gctccgagca tccggtcgca 540
ggcacggggc tcggcttgctc gatcagccag cgcctggcgc agctaataaa tggcagtcgt 600
aaactgggtc gtgagctggg gttgggcagc agcttttagc tcagggttcc gcttgagcgg 660
atcgcgatgc aggctgagcc gcaggaccta gccgggtgcg ccgtccaagt gctggcgcc 720
gtccgcgacc taacggaatg cctgtgtggc tggatctccc gctgggggtg aagggccatg 780
gtcgcgacgc cgaggctcgt ggacgaggcg gacgcgacct cgctgctggg cgaagtgtta 840
ctgctggagg gggcgccgat gttcgaagca tggccaggat gccgggtgga gctttccct 900
cagggtgata tggagccgca ggcacagggc cgcgactggc tgctcgggct caacaacctg 960
gacggcctgc atcgtgctct gggcctggcc catgggcgtc tcgctgatcc ttcgacgcgc 1020
ccgatacgcc tggctccgtt gcgcaatcta ggtctccgcg tcctagtggg ggaggataac 1080
gcgatcaacc agttgatctt gagggaccag atggaagcgc tgggctgcag cgtggagctg 1140
ctcttcgatg gtcgcgaggg gttgctgcac tgccagacgg cctgcttcga cgtggtgctc 1200
accgatatca acatgccgaa catgaacgga tacgagctaa ccgcggagct acggcgccaa 1260
gggttccggc agccgatcat cggcgcgacg gcgaacgcca tgcgtgagga gcgcgagcgc 1320
tgcatgtccg ccgggatgaa cgattgcctg gtcaaaccgg tggatctgaa tgcccttcag 1380
aactgcttga ttaatatctt caaggtggat cgatga 1416

```

<210> 4

<211> 471

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 4

```

Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu
1      5      10      15
Glu Leu Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu
20     25     30
Lys Ala Ile Gln His Ser Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp
35     40     45
Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys
50     55     60
Val Glu Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe
65     70     75     80
Thr Gly Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser
85     90     95
Ala Glu Leu Pro Leu Arg Met Arg Gly Ala Ala Ala Ser Ile Arg Gln
100    105    110
Ile Leu Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly
115    120    125
Tyr Val Asn Val His Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val
130    135    140
Met Leu Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu
145    150    155    160
Asp Gln Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu
165    170    175
His Pro Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu
180    185    190
Ala Gln Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu
195    200    205
Gly Ser Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln
210    215    220
Ala Glu Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro
225    230    235    240
Val Arg Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly

```

<400>	5								
atgatggatg	ttatacggga	gc atgaggt a	t t t c t t g g g c	g c a t c g c t c g	a a a a g c g a c	60			
aagaccaccc	agaagtacga	c t a t g a c g t g	g t g c c t t t g c	a g c g g c a c t b	g t t g g c a a a g	120			
gaaaacggat	tagcggtcta	t g a g g a c g g g	g a g t t t t c o c t	t t g c t a t g c c	a t t t c t a c t g	180			
gctaccaagc	acgcgttgag	c g c c g a t t c c	t c g g g a g a t c	c g t t t t c g c t	c g g t g t a t t g	240			
ctcgccaatt	tctacggaag	c t t c t g g a g t	g t t t c c g c c t	a t c c c g c g c c	a c a g t t a c t g	300			
atctttgatc	tttcgggcag	c a c c c g c c t g	g c a g t g c c g t	c g a t t c c c t c	c a c a g c g c a g	360			
cgtgacagg	tgagcggaa	c t a t c c g a t c	a t a g t c g a g c	g c a t t c t g g c	g c g c t t g c g c	420			
accggccggg	tgggggaagg	c g c t c a g c g t	g t c a t t t g g a	t a c g c g t g c	t g c t a t c g c	480			
gactcggcgc	tggagatgtt	g g g a g t c g c c	c g g g t t g a t c	t g c c g g a a a c	a c t c t g g t g g	540			
cacgacgagc	cgaaccatct	g a t c a t c g c t	g c g a g c c t g c	t t g a t c t c a g	g c g a a t c a a t	600			
gacttcgaac	agttgggtga	g c g c c c g g c a	t t c g a t t c g t	a c a g c c t g g t	a t c g c c g g a t	660			
ggcgagggtat	tgctgcgtca	g g c c c c t g c g	a c c g g c c t g a	g g g a t g g c c t	g a a c t c a a c	720			
cgacaggggg	tcgccgttgc	a c t g c g c a g c	c a g c c t g a g a	a c g g c t g g c t	c g c g g t c t a c	780			
cgaaccgact	acggcaattt	c t t t c g c c a c	t c c c g g t g g c	t g g t g g c a g g	t c t g c t g c t g	840			
accccggcgc	tgctcctggc	c g g t t g g t c t	g g g a t g c g t t	g g t a c a c c a g	c a g c g t c g t c	900			
aacctgggtg	atcgggcgca	c c g g c a a c t g	g t g g a g a g c g	a c a c c t t c a g	c g g a c g c t g	960			
atacacgacg	cgccgggtgc	t c t g g t g g t g	c t g a c c c a g a	a t g a c c a g c a	a c t g g t g a c c	1020			
tgcaaccact	tggccgcccc	g t g g c t g g g c	g g g c c c a c g g	a g a t c c t t g g	g c t g a c t t c c	1080			
aactggaagc	tttcagatgc	g c g t g g g c a g	g t a c a c g a g g	a c a t c t g t a t	c c a g g t c g g t	1140			
gggcgctatt	tggacacgcg	c t t c g c g g c g	a c c c g c t a t g	c c g g c a c c g a	g g c g g t a c t g	1200			
tgctatttca	accacatcac	q q t c c a c t g c	g a g g c g g a g a	c c g c g c t g t c	c a a t g c g a a g	1260			

```

cgagcagcgg atgccgccag ccaggccaag accctgttcc tggcccgcat gagccatgaa 1320
atccgtactc ccctgtacgg tgtccttggc accctggagt tgctcgacct gaccaccctg 1380
aacgagcggc aacgcgccta cctacgcacc atccagagtt cgtctgcgac gctcatgcaa 1440
ctgattagcg atgtgctgga tgtctogaag atcgaagcgg ggcagatggc tctgaccctg 1500
gccgccttca atccgctgga cctagtgcgg gaagtgcctt gcaactttgc cgccagcgcc 1560
atggccaagg acctgcaggt agaccgcgtc gatactcttg cgcttgaggc gcaggctcg 1620
catggcttcg aagaaagcgt tctgttcgag gttgctgggt gctcggtcgg ccatttcgaa 1680
gaggggtgctg tcggcggtgt cgaacaacgc ctgcaacgcc tgtttcagct gcagcgccgc 1740
cttgtcgcgc acctgcacga ggatgaccgg caggcgcccc gctccggcgt tcggcgacgg 1800
ctcgggaagcg accctggtca ggtgcaccac attggcatcg ttctgcatcg ggactctcct 1860
gccaccctcg cggccgcgca tggaatggca aaaatcgggc acagaggatc gattggcgtc 1920
gtccgtaacg tcaatttcca ggcgtaaaaa acaagtatct acattcatta tagagatact 1980
ttcaaatacta gatag 1995

```

<210> 6

<211> 664

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 6

```

Met Met Asp Val Ile Arg Glu His Glu Val Phe Leu Gly Arg Ile Ala
1          5          10          15
Arg Lys Ser Asp Lys Thr Thr Gln Lys Tyr Asp Tyr Asp Val Val Pro
20          25          30
Leu Gln Arg His Leu Leu Ala Lys Glu Asn Gly Leu Ala Val Tyr Glu
35          40          45
Gly Arg Glu Phe Ser Phe Ala Met Pro Phe Leu Leu Ala Thr Lys His
50          55          60
Ala Leu Ser Ala Asp Ser Ser Gly Asp Pro Phe Ser Leu Gly Val Leu
65          70          75          80
Leu Ala Asn Phe Tyr Gly Ser Phe Trp Ser Val Ser Ala Tyr Pro Ala
85          90          95
Pro Gln Leu Leu Ile Phe Asp Leu Ser Gly Ser Thr Arg Leu Ala Val
100          105          110
Pro Ser Ile Pro Ser Thr Ala Gln Arg Asp Arg Leu Ser Gly Ser Tyr
115          120          125
Pro Met Ile Val Glu Arg Ile Leu Ala Arg Leu Arg Thr Arg Pro Val
130          135          140
Gly Glu Asp Ala Gln Arg Val His Trp Ile Arg Ala Asp Arg Tyr Arg
145          150          155          160
Asp Ser Ala Leu Glu Met Leu Gly Val Ala Arg Val Asp Leu Pro Glu
165          170          175
Thr Leu Trp Trp His Asp Glu Pro Asn His Leu Ile Ile Ala Ala Ser
180          185          190
Leu Leu Asp Leu Arg Arg Ile Asn Asp Phe Glu Gln Leu Val Glu Arg
195          200          205
Pro Ala Phe Asp Ser Tyr Ser Leu Val Ser Pro Asp Gly Glu Val Leu
210          215          220
Leu Gly Ala Ala Pro Ala Thr Gly Leu Arg Asp Gly Leu Asn Leu Thr
225          230          235          240
Arg Gln Gly Val Ala Val Gln Leu Arg Ser Gln Pro Glu Asn Gly Trp
245          250          255
Leu Ala Val Tyr Arg Thr Asp Tyr Gly Asn Phe Phe Arg His Ser Arg
260          265          270
Trp Leu Val Ala Gly Leu Leu Leu Thr Pro Ala Leu Leu Ala Gly
275          280          285
Trp Leu Gly Met Arg Trp Tyr Thr Ser Ser Val Val Asn Pro Val His
290          295          300

```

Arg Ala His Arg Gln Leu Val Glu Ser Asp Thr Phe Ser Arg Thr Leu
 305 310 315 320
 Ile Gln Thr Ala Pro Val Ala Leu Val Val Leu Thr Gln Asp Asp Gln
 325 330 335
 Gln Leu Val Thr Cys Asn His Leu Ala Ala Gln Trp Leu Gly Gly Pro
 340 345 350
 Thr Glu Ile Leu Gly Leu Thr Ser Asn Trp Lys Leu Phe Asp Ala Arg
 355 360 365
 Gly Gln Val Pro Gly Asp Ile Cys Ile Gln Val Gly Gly Arg Tyr Leu
 370 375 380
 Gln Thr Ala Phe Ala Ala Thr Arg Tyr Ala Gly Thr Glu Ala Val Leu
 385 390 395 400
 Cys Val Phe Asn Asp Ile Thr Val His Cys Glu Ala Glu Thr Ala Leu
 405 410 415
 Ser Asn Ala Lys Arg Ala Ala Asp Ala Ala Ser Gln Ala Lys Thr Leu
 420 425 430
 Phe Leu Ala Arg Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Val
 435 440 445
 Leu Gly Thr Leu Glu Leu Leu Asp Leu Thr Thr Leu Asn Glu Arg Gln
 450 455 460
 Arg Ala Tyr Leu Arg Thr Ile Gln Ser Ser Ser Ala Thr Leu Met Gln
 465 470 475 480
 Leu Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Met
 485 490 495
 Ala Leu Thr Leu Ala Ala Phe Asn Pro Leu Asp Leu Val Arg Glu Val
 500 505 510
 Leu Gly Asn Phe Ala Ala Ser Ala Met Ala Lys Asp Leu Gln Val Asp
 515 520 525
 Pro Leu Asp Thr Leu Ala Leu Glu Ala Gln Val Ala His Gly Phe Glu
 530 535 540
 Glu Ser Val Leu Phe Glu Val Ala Gly Gly Ser Val Gly His Phe Glu
 545 550 555 560
 Glu Gly Val Val Gly Val Val Glu Gln Arg Leu Gln Arg Leu Phe Gln
 565 570 575
 Leu Gln Arg Arg Leu Val Ala His Leu His Glu Asp Asp Arg Gln Ala
 580 585 590
 Pro Arg Ser Gly Val Arg Arg Arg Leu Gly Ser Asp Pro Gly Gln Val
 595 600 605
 His His Ile Gly Ile Val Leu His Arg Asp Ser Pro Ala Thr Leu Ala
 610 615 620
 Ala Ala His Gly Met Ala Lys Ile Gly His Arg Gly Ser Ile Gly Val
 625 630 635 640
 Val Arg Asn Val Asn Phe Gln Ala Ser Lys Thr Ser Ile Tyr Ile His
 645 650 655
 Tyr Arg Asp Thr Phe Lys Ser Arg
 660

<210> 7

<211> 6410

<212> DNA

<213> Pseudomonas aeruginosa PA14

<400> 7

cttcctcacc gcactttcat tcaccgcacc gttatcgggc tcatggacaa tgcctcgcag 60
 tcctgagacc tcattaggaa tcattctgata gaggggaattt ttccatgccg tcagtccttc 120
 ggcatgaagt tgaagaattt cttacagcct ttgatagcg gtttctccac tccgagtgct 180
 gcgctcaagc tgctccgcac gctcggtggc gccttgatgt tgtgcgtgct atgcagcctg 240

atattcagtg	tgagcatggg	tttaaaccat	cagggtgtccc	tcagtcggca	agctatgaat	300
gtggctatgt	acgaagcgca	gctttatttc	gagcagcgcg	aggcgttgct	caatcacttg	360
agcggcaatg	tcgtgccctt	ggccgcgggt	agagcgctcg	tcaacgaagc	gccgaacaat	420
gtgagcatcc	tgccgttgag	tgacggaggg	cgaggctctgc	tattgaccgc	tcgcacgctc	480
ggtgatctcc	gggaaaagcg	gctggcactg	atgtatctgg	tcgataccga	caaaggccct	540
ctggtttacc	ggcttaccgc	cgatggtagg	ccctcggcag	cgatatccag	cacgataacc	600
aaagagggtg	accgagcctt	gctggcgact	ccgtcggcgc	ctgttcactg	ggtgactgac	660
ggtggtaccc	ctcaacggct	gtaccttttt	gaatccttag	gcgatgagcc	gggagggggg	720
tggttagggc	tggagattct	cggcgaagac	ctcgattcga	tggtgcgcgc	gaatgatgcc	780
ggaaactaca	tgctgctgga	tcagcatggg	caggctcgta	tcgctacgga	cgcagaggcg	840
ctggggagcg	gtgcgtcgcg	gacgcttttg	cgtggagacg	gcttcgggtt	catcgggtgc	900
ggccacactg	cgagcatat	gggtgctttt	cagcacgtgg	ggctcttcgag	ctgggatctg	960
atctatcaca	tcggtatcgg	tcgcctgttg	ctggctctgt	ggctccctct	gttacttgcc	1020
tctgcgttgg	cactcgagct	cggcatccta	ctgcattggc	tggtgcggag	catcgagcga	1080
cgcttgatag	agcccgcaaa	gcgacgcctt	gaagcattga	aggagagcga	agcctttttc	1140
cgtgcagtta	tccaggccgc	gcccgtcgcg	ctgtgcgtgc	tgcgctcgtc	cgacgccgca	1200
gtggtcctgg	aaaatcccca	ggcgcgccaa	tggtcgggtg	atagcgaggg	gattgcccac	1260
gacgcgccga	gatggatttc	ccaggcgctt	gcaggagggtg	tgaagtgttc	tggagaagaa	1320
ctggaaaaccg	aggcagggct	acatcttcat	ctcaattaca	cgcccaccgc	ctataacggt	1380
gaagacgtat	tgttctgcgc	cttcagtgaa	atcagtgac	gcaagcggat	ggaggcgga	1440
ctggctcgcg	caaaatccct	ggcggatgct	gccaatgaag	ccaagacgct	gtttctcgcc	1500
accatgagcc	atgaaatccg	cacacctctg	tacggcatgc	ttggcacgct	tgagctgctt	1560
gggcttaccg	agctgagtcg	gcagcaggcc	ggttaccta	aggcaatcca	gcattcctcg	1620
tcgacctcgc	tgcaactgat	cagcgatgtg	cttgacgtat	ccaagataga	ggccggccaa	1680
ctggactcag	agtcgctgga	attctcccg	ctgggaattga	ccgaagaggt	cgtgcagtcg	1740
ttcacccgtg	ccgcgcaggc	caaggggctg	cagttgtata	cctgcctctc	tgccgagctg	1800
ccgctgcgca	tgccggggggc	cgccggcgctg	atccggcaga	ttctcaacaa	cctgctgagc	1860
aacgcggtga	agttcaccca	caatggctat	gtcaacgtcc	acctgaaggc	cagcgtgggtc	1920
gatgcgcaat	gtgtgatgct	gacctggcag	gtcaacgata	ccggcatggg	gatcaacgctc	1980
gaggatcagc	cgcgtctgtt	cgaaccgttc	taccagatac	gcccgtccga	gcateccggtc	2040
gcaggcacgg	gcctcggctt	gtcgatcagc	cagcgctctg	cgcagcta	gaatggcagt	2100
ctgaaactgg	tcagtgagct	ggggttgggc	agcagcttta	gcctcaggct	tcgcgttgag	2160
cggatcgcca	tgcaaggctga	gccgcaggac	ctagccgggt	gcgcggtcca	agtgtgtggc	2220
cctgtccgcg	acctaaccga	atgcctgtgt	ggctggatct	cccgtggggg	tgggaagggcc	2280
atggtcgcga	cgccgagggtc	gctggacgag	gcggacgcga	cctcgctgct	ggtcgaagtg	2340
ttactgctgg	agggggcgcc	gatgttcgaa	gcatggccag	gatgccgggt	ggagctttcc	2400
cctcagggtg	atatggagcc	gcaggcacag	ggccgcgact	ggctgctcgg	gctcaacaac	2460
ctggagggcc	tgctcgctgc	tctgggcctg	gcccattggc	gtctcgctga	tccttcgacg	2520
ccgcccatac	ggctggctcc	gttgcgcaat	ctaggtctcc	gcgtcctagt	ggtggaggat	2580
aacgcgatca	accagttgat	cttgagggac	cagatggaag	cgctgggctg	cagcgtggag	2640
ctgctcttcg	atggctcgca	ggcgttgctg	cactgccaga	cggcctgctt	cgcagtggtg	2700
ctcacccgata	tcaacatgcc	gaacatgaac	ggatacagac	taaccgcgga	gctacggcgc	2760
caagggttcc	ggcagccgat	catcggcgcg	acggcgaaacg	ccatgcgtga	ggagcgcgag	2820
cgctgcatgt	ccgcggggat	gaacgattgc	ctggtcaaac	cggtggatct	gaatgcctt	2880
cagaactgct	tgattaatat	tctcaagggtg	gatcgatgag	ctggaaatcc	tatcgggtgc	2940
tggtggtcga	agatcagccg	tttcagcgcg	aatacctgct	caacctgttt	cgcgagcgcg	3000
gcgtgcagta	cctggtaggt	gccggcgacg	gcgcggaggc	gttgcgctgc	ctgaagcagg	3060
acaggttcga	cctgatcctc	agcgatctga	tgatgccggg	catggatggt	atccaaatga	3120
tcctgcaact	gccgtatctc	aagcatcgtc	cgaagctggc	gctgatgagc	tcctcgctgc	3180
agcggatgat	gctcagtgcc	agccgggtcg	cccagagtct	cggtctgtcg	gtaatcgacc	3240
tgttgcccaa	gccgactctg	cccaaggcca	ctggccaaact	tctggaacac	ctggaaagat	3300
gcctcaggca	gaagctggag	ccggaaaaccg	acgagactcc	gcatggggcg	acggcggtgc	3360
tggatgccct	gcataacgag	caactggtga	cctggttcca	ggctaagaaa	tcctccaca	3420
ccgggcgcac	agtccggcgcc	gaggcggtga	tacgctggag	ccacccgcag	catggcctgt	3480
tgctgccacg	ctgtttcatg	agtgatgtcg	acgctaccgg	tctgcacgag	gcgttgctct	3540
ggcgcgtgct	cgaacagacc	ctgaacgccc	aggaatcgctg	gcgcagggcg	ggttacgaga	3600
ttccgggttc	ggtgaatctg	ccgcgcgacc	tgctcgataa	ccaggaactt	ccggatcgac	3660
tctatgagta	cgtcggcgct	cgcggggctt	gtaccagctc	actatgtttc	gagttgaccg	3720

```

agagcagtggt cacaactctg tcaagtaact actatgcagg tgcctgtcgc ttgcgcatga 3780
aagggttcgg attggcccag gacgactttg gccagggtta cagctcgttc tataacctgg 3840
tcacgacgcc tttcacggag ctgaagatcg accgctccct agtccaggga tgcgtagagg 3900
ataacggcct caatgcagct gtcacagtt gtattgagtt gggtcaccgc ctgaatctcg 3960
acgtggtggc cgaaggcgtg gagacctgcg aggaactgaa tcttcttcgt cgtcttggct 4020
gcgaccgggc gcagggtttc ctgatttcta aggcagtgtc tgctcgtgag ttcgagcggc 4080
agttaaggga ggacggcccc agcctccttg tttaacgcag tatccccatt atcgcgaggt 4140
cgatcgagc caaccaccgt cagcgcaaca gtgtcgtga aggaggccgc atcccgta 4200
gtctgtagt gccttggagc acgacaacaa acttttgctc aaatggacaa ccctctcgca 4260
gagctgagc atcggttga tctgtgtggt ggtgctgacc gtattgctgt tcagcatctg 4320
ttctgtgtcg ctggggagat tgtttcagga ggaggaggac aaagtctcct tccacttcac 4380
ccgtatgatg gatgttatac gggagcatga ggtatttctt gggcgcatcg ctgaaaaag 4440
cgacaagacc acccagaagt acgactatga cgtggtgcct ttgcagcggc acttggtggc 4500
aaaggaacac ggattagcgg tctatgaggg acgggagttt tcctttgcta tgccatttct 4560
actggctacc aagcacgcgt tgagcgccga ttctcggga gatccgtttt cgctcgggtg 4620
attgtcgcc aatttctacg gaagcttctg gagtgttcc gcctatccc cgccacagtt 4680
actgatcttt gatctttccg gcagcaccgc cctggcagtg ccgtcgattc cctccacagc 4740
gcagcgtgac aggttgagcg gaagctatcc gatgatagtc gagcgcatc tggcgcgctt 4800
gcgcaccggc ccggtggggg aggacgctca gcgtgtccat tggatacgcg ctgatcgcta 4860
tcgcgactcg gcgctggaga tgttgggagt cgcccggtt gatctgcgg aaacactctg 4920
gtggcacgac gagccgaacc atctgatcat cgctcgagc ctgcttgatc tcaggcgaat 4980
caatgacttc gaacagttgg ttgagcggcc ggcattcgat tcgtacagcc tggtatcgcc 5040
ggatggcgag gtattgctcg gcgcggcccc tgcgaccggc ctgagggatg gccgaacct 5100
caccgcagac ggggtcgccg ttcaactgcg cagccagcct gagaacggct ggctcgcggt 5160
ctaccgaacc gactacggca atttctttcg ccactccgg tggtgtgtgg caggctgtgt 5220
gctgaccccg gcgctgctcc tggccggttg gctcgggatg cgttggtaca ccagcagcgt 5280
cgtcaacccg gtgcatcggg cgacccggca actggtggag agcgacacct tcagccggac 5340
gctgatacag accgcgcggc tggctctggt ggtgctgacc caggatgacc agcaactggt 5400
gacctgaac cacttggccg ccagtggtg gggcgggccc acggagatcc ttgggctgac 5460
ttccaaactgg aagcttttcg atgcgcgtgg gcaggtacca ggagacatct gtatccaggt 5520
cgggtgggccc tatttgcaga ccgccttcgc ggcgacccgc tatgccggca ccgaggcggt 5580
actgtgcgta ttcaacgaca tcacggtcca ctgcgaggcg gagaccgcgc tgtccaatgc 5640
gaagcgagca gcggatgccc ccagccagc caagaccctg ttcttggccc gcatgagcca 5700
tgaaatccgt actcccctgt acggtgtcct tggcaccctg gagttgctcg acctgaccac 5760
cctgaacgag cggcaacgcg cctacctacg caccatccag agttcgtctg cgacgctcat 5820
gcaactgatt agcgatgtgc tggatgtctc gaagatcgaa gcggggcaga tggctctgac 5880
cctggccggc ttcaatccgc tggacctagt gcgggaagtg cttggcaact ttgccgccag 5940
cgccatggcc aaggacctgc aggtagacct gctcgatact cttgcgcttg aggcgcaggt 6000
cgcgcatggc ttogaagaaa gcgttctggt cgaggttgct ggtggctcgg tcggccattt 6060
cgaagagggg gtctgctggcg ttgtcgaaca acgcctgcaa cgctgttttc agctgcagcg 6120
ccgccttgtc gcgcacctgc acgaggatga ccggcaggcg ccccgctccg gcgttcggcg 6180
acggctcgga agcgacctg gtcagggtgca ccacattggc atcgttctgc atcgggactc 6240
tctgccacc ctgcggccg cgcattggaat ggcaaaaatc gggcacagag gatcgattgg 6300
cgctgctcgt aacgtcaatt tccaggcgtc aaaaacaagt atctacattc attatagaga 6360
tactttcaaa tctagataga gggtttctct atagacatga ctctcttcac 6410

```

<210> 8

<211> 2796

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 8

```

atgaagttga agaatttctt acagcctttt gatagcgggt tctccactcc gagtgtcgcg 60
ctcaagctgc tccgcatgct cgggtggcgcc ttgatgttgt gcgtgctatg cagcctgata 120
ttcagttgta gcatgggttt aaaccatcag gtgtccctca gtcggcaagc tatgaatgtg 180
gctatgtacg aagcgcagct ttatttctcag cagcgcgagg cgttgctcaa tcacttgagc 240

```

```

ggcaatgtcg tgccttggc cgcgggtaga gcgctcgtca acgaagcgcc gaacaatgtg 300
agcatcctgc cgttagtgga cggagggcga ggtctgctat tgaccgctcg caccgctcgt 360
gatctccggg aaaagcggct ggcactgatg tatctggctg ataccgacaa aggccctctg 420
gtttaccggc ttaccgccga tggtagggcc tcggcagcga tatccagcac gataaccaa 480
gaggtgtacc gagccttgct ggcgactccg tcggcgcctg ttcactgggt gactgacggt 540
ggtacccctc aacggctgta cctttttgaa tccttaggcg atgagccggg cgaggggtgg 600
ctaggcctgg agattctcgg cgaagacctc gattcgatgt tgcgccgaa tgatgccgga 660
aactacatgc tgctggatca gcatgggcag gtcgtactcg ctacggacgc agaggcgtg 720
gggagcgggtg cgtcgcggac gctttttgct ggagacggct tcggtttcat cgggtgctggc 780
ccactgccgc agcatatggt gctttttccag cacgtggggg cttcgagctg ggatctgac 840
tatcacatcg gtatcggtcg cctgttgctg gctctgtggc tccctctgtt acttgccctc 900
gcgttgccac tcgcagtcgg cactctactg cattggctgg tgcggagcat cgagcgacgc 960
ttgatagagc ccgcaaagcg acgccttgaa gcattgaagg agagcgaagc cttttccctg 1020
gcagttatcc aggcgcgcgc cgtcgcgctg tgcgtgctgc gtcgtgccga cgccgcagtg 1080
gtcctggaaa atccccaggc gcgccaatgg ctgggtgata gcgagcgcat tgcccacgac 1140
gcgcgagat ggatttccca ggcgttcgca ggaggtgtga agtggtctgg agaagaactg 1200
gaaaccgagg cagggtaca tcttcatctc aattacacgc ccacccgcta taacggtgaa 1260
gacgtattgt tctgcgcctt cagtgaatc agtgacgca agcggatgga ggccgaaactg 1320
gctcgcgcaa aatccctggc ggatgctgcc aatgaagcca agacgctgtt tctcgccacc 1380
atgagccatg aaatccgcac acctctgtac gcacgcttga gctgctggg 1440
cgtaccgagc tgagtcggca gcaggccggg tacctaaagg caatccagca ttcctcgtcg 1500
accctgctgc aactgatcag cgatgtgctt gacgtatcca agatagaggc cggccaaactg 1560
gacctagagt gcgtggaatt ctccccgctg gaattgaccg aagaggctcg gcagtcgttc 1620
accggtgccg cgcaggccaa ggggctgcag ttgtatacct gcctctctgc ggagctgccg 1680
ctgcgcatgc gggggggcgc ggctcgatc cggcagattc tcaacaacct gctgagcaac 1740
gcggtgaagt tcaccgacaa tggctatgtc aacgtccacc tgaaggccag cgtggtcgat 1800
gccgaatgtg tgatgctgac ctggcaggtc aacgataccg gcatggggat caacgtcgag 1860
gatcagccgc gtctgttcga accgttctac cagatacgcc gctccgagca tccggctcgca 1920
ggcacgggcc tcggcttgct gatcagccag cgcttgccgc agctaattgaa tggcagctcg 1980
aaactggtca gtgagctggg gttgggcagc agctttagcc tcaggcttcc gcttgagcgg 2040
atcgcgatgc aggtgagcc gcaggaccta gccgggtgcg ccgtccaaagt gctggcgcc 2100
gtccgcgacc taacggaatg cctgtgtggc tggatctccc gctggggtgg aagggccatg 2160
gtcgcgacgc cgaggtcgct ggacgagggc gacgcgacct cgctgctggt cgaagtgtta 2220
ctgctggagg gggcgccgat gttcgaagca tggccaggat gccgggtgga gctttccct 2280
cagggtgata tggagccgca ggcaaggcc cgcgactggc tgctcgggct caacaacctg 2340
gacggcctgc atcgtgctct gggcctggcc catgggcgtc tcgctgatcc ttcgacgccg 2400
ccgatacgcc tggctccgtt gcgcaatcta ggtctccgcg tcctagtggg ggaggataac 2460
gcgatcaacc agttgatctt gagggaccag atggaagcgc tgggctgcag cgtggagctg 2520
ctcttcgatg gtgcgagggc gttgctgcac tgccagacgg cctgcttcga cgtggtgctc 2580
accgatatca acatgccgaa catgaacgga tacgagctaa ccgcggagct acggcgccaa 2640
gggttccggc agccgatcat cggcgcgacg gcgaacgcca tgcgtgagga gcgcgagcgc 2700
tgcatgtccg ccgggatgaa cgattgcctg gtcaaacggg tggatctgaa tgcccttcag 2760
aactgcttga ttaattattc caaggtggat cgatga 2796

```

<210> 9

<211> 2721

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 9

```

atgctcggtg ggccttgat gttgtgcgtg ctatgcagcc tgatattcag tgtgagcatg 60
gttttaaac atcaggtgtc cctcagtcgg caagctatga atgtggctat gtacgaagcg 120
cagctttatt tcgagcagcg cgaggcgttg ctcaatcact tgagcggcaa tgcgtgcc 180
ttggccgcgg gtagagcgtc cgtcaacgaa gcgccgaaca atgtgagcat cctgccgttg 240
agtgcgggag ggcgaggtct gctattgacc gctcgcacgc tcggtgatct ccgggaaaag 300
cggctggcac tgatgtatct ggtcgatacc gacaaaggcc ctctggttta ccgcttacc 360
gccgatggta ggccctcggc agcgatatcc agcacgataa ccaaagaggt gtaccgagcc 420
ttgctggcga ctccgtcggc gcctgttcac tgggtgactg acgggtgtac ccctcaacgg 480

```

```

ctgtaccttt ttgaatcctt aggcgatgag cggggcgagg ggtggctagg cctggagatt 540
ctcggcgaag acctcgattc gatgttgccg cggaaatgat cctggtgctg 600
gatcagcatg ggcaggtcgt actcgctacg gacgcagagg cgtggggag cgggtcgtcg 660
cggaacgttt tgcgtggaga cggcttcggt ttcacggtg ctggcccaat gccgcagcat 720
atgggtgcttt tccagcacgt ggggtcttcg agctgggatc tgatctatca categgtatc 780
ggtcgccctgt tgcgtgctct gtggctccct ctgttacttg cctctgcgtt ggcaactcga 840
gtcggcatcc tactgcattg gctgggtcgg agcatcgagc gacgcttgat agagcccgca 900
aagcgacgcc ttgaagcatt gaaggagagc gaagcctttt cccgtgcagt tatccaggcc 960
gcgcccgtcg cgtgtgctgt gctgcgtcgt gccgacgcgc cagtggctct ggaaaaatcc 1020
caggcgcgcc aatggctggg tgatagcgag gcgattgccc acgacgcgcc gagatggatt 1080
tcccaggcgt tcgcaggagg tgtgaagtgt tctggagaag aactggaaac cgaggcaggg 1140
ctacatcttc atctcaatta cagccccacc cgctataacg gtgaagacgt attgttctgc 1200
gccttcagtg aaatcagtg acgcaagcgg atggaggcgg aactggctcg cgcaaaatcc 1260
ctggcgggatg ctgccaatga agccaagacg ctgtttctcg ccaccatgag ccatgaaatc 1320
cgcacacctc tgtacggcat gcttggcacg cttgagctgc ttgggcgtac cgagtgcagt 1380
cggcagcagg ccggttacct aaaggcaatc cagcatctct cgtcgaccct gctgcaactg 1440
atcagcgatg tgcttgacgt atccaagata gaggccggcc aactggacct agagtgcgtg 1500
gaattctccc cgttggaatt gacogaagag gtcgtgcagt cgttcaccgg tgccgcgcag 1560
gccaaagggc tgcatgtgta tacctgcctc tctgcggagc tgccgctgcg catgccccgg 1620
gccgcggcgt gatccggca gattctcaac aacctgctga gcaacgcggg gaagttcacc 1680
gacaatggct atgtcaacgt ccacctgaag gccagcgtgg tcgatgccga atgtgtgatg 1740
ctgacctggc aggtcaacga tacggcgatg gggatcaacg tcgaggatca gccgcgtctg 1800
ttcgaaacgt tctaccagat acgcccgtcc gacgatccgg tcgcaggcac gggcctcggc 1860
ttgtcgatca gccagcgcct ggcgcagcta atgaatggca gtctgaaact ggtcagtgcg 1920
ctggggttgg gcagcagctt tagcctcagg ctcccgcttg agcggatcgc gatgcaggct 1980
gagccgcagg acctagccgg gtgcgcctc caagtgcctg cgcctgtccg cgacctaacg 2040
gaatgcctgt gtggctggat ctcccgtctg ggtggaaggg ccatggtcgc gacgccgagg 2100
tcgctggacg aggcggacgc gacctcgctg ctggtcgaag tggtactgct ggagggggcg 2160
ccgatgttcg aagcatggcc aggatgccgg gtggagcttt cccctcaggg tgatattggg 2220
ccgcaggcac agggccgcga ctggctgctc gggctcaaca acctggacgg cctgcacgtg 2280
gctctgggcc tggcccatgg gcgtctcgct gatccttoga cgccgccgat acggctggct 2340
ccgttcgcga atctaggtct ccgctccta gtggtggagg ataacgcgat caaccagttg 2400
atcttgaggg accagatgga agcgtggggc tgcagcgtgg agctgctctt cgatggtcgc 2460
gaggcgctgc tgcactgcca gacggcctgc ttcgacgtgg tgctcaccga tatcaacatg 2520
ccgaacatga acggatacga gctaaccgcg gagctacggc gccaaagggt ccggcagccg 2580
atcatcgccg cgacggcgaa cgccatgcgt gaggagcgcg agcgtgcgt gtccgcggcg 2640
atgaacgatt gcctggtcaa accggtggat ctgaatgccc ttcagaactg cttgattaat 2700
attctcaagg tggatcgatg a

```

<210> 10

<211> 2703

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 10

```

atgttgtgctg tgctatgcag cctgatattc agtgtgagca tggtttttaa ccatcagggtg 60
tccctcagtc ggcaagctat gaatgtggct atgtacgaag cgcagcttta tttcgagcag 120
cgcgaggcgt tgctcaatca cttgagcggc aatgtcgtgc ccttggccgc gggtagagcg 180
ctcgtcaacg aagcgccgaa caatgtgagc atcctgcctg tgagtacggg agggcgagggt 240
ctgctattga ccgctcgcac gctcgggtgat ctccgggaaa agcggctggc actgatgtat 300
ctggctcgata ccgacaaagg cctctggtt taaccggcta ccttgctggc taggcctcg 360
gcagcgatat ccagcacgat aaccaaagag gtgtaccgag ccttgctggc gactccgtcg 420
gcgcctgttc actgggtgac tgacgggtgg acccctcaac ggctgtacct ttttgaatcc 480
ttaggcgatg agcggggcga ggggtggcta ggccctggaga ttctcggcga agacctcgat 540
tcgatgttgc gccggaatga tgccggaaac tacatgctgc tggatcagca tgggcaggctc 600
gtactcgcta cggacgcaga ggcgtgggg agcgggtcgt tttgcgtgga tttgcgtgga 660
gacggcttcg gtttcatcgg tgctggccca ctgcccgcagc atatgggtgct tttccagcac 720
gtggggctct cgagctggga tctgatctat cacatcggta tcggctgcct gttgctggct 780

```



```

ctgtggctcc ctctgttact tgcctctgcg ttggcactcg cagtcggcat cctactgcat 840
tggctggtgc ggagcatcga gcgacgcttg atagagcccg caaagcgacg ccttgaagca 900
ttgaaggaga gcgaagcctt ttcccggtga gttatccagg ccgcgcccgt cgcgctgtgc 960
gtgctgcgtc gtgccgacgc cgcagtggtc ctggaaaatc cccaggcgcg ccaatggctg 1020
ggtgatagcg aggcgattgc ccacgacgcg ccgagatgga tttcccaggc gttcgcagga 1080
ggtgtgaagt gttctggaga agaactggaa accgaggcag ggctacatct tcatctcaat 1140
tacacgocca cccgctataa cgggtgaagac gtattgttct gcgccttcag tgaaatcagt 1200
gcacgcaagc ggatggaggc ggaactggct cgcgcaaaat ccctggcgga tgctgccaat 1260
gaagccaaga cgctgtttct cgccaccatg agccatgaaa tccgcacacc tctgtacggc 1320
atgcttggca cgcttgagct gcttggggct accgagctga gtcggcagca ggccggttac 1380
ctaaaggcaa tccagcattc ctgctcgacc ctgctgcaac tgatcagcga tgtgcttgac 1440
gtatccaaaga tagaggccgg ccaactggac ctagagtgcg tggaattctc cccgctggaa 1500
ttgaccgaag aggtcgtgca gtcgttcacc ggtgcgcgcg aggccaaagg gctgcagttg 1560
tatacctgcc tctctgcgga gctgcgcgtg cgcgtgcggg gggccgcggc gtcgatccgg 1620
cagattctca acaacctgct gagcaacgcg gtgaagttca ccgacaatgg ctatgtcaac 1680
gtccacctga aggccagcgt ggtcgatgcc gaatgtgtga tgctgacctg gcaggtcaac 1740
gataccggca tggggatcaa cgtcgaggat cagccgcgtc tgttcgaacc gttctaccag 1800
atacgcgct cccgacatcc ggtcgcaggc acgggcctcg gcttgtcgat cagccagcgc 1860
ctggcgacgc taatgaatgg cagtctgaaa ctggtcagtg agctggggtt gggcagcagc 1920
tttagcctca ggcttccgct tgagcggatc gcgatgcagg ctgagccgca ggacctagcc 1980
gggtgcgcgc tccaagtgcg ggcgccctgtc cgcgacctaa cggaatgcct gttgtggtgg 2040
atctcccgct ggggtggaag ggccatggtc gcgacgcgga ggtcgcgtgga cgaggcgga 2100
gcgacctcgc tgctggtcga agtggtactg ctggaggggg cgccgatggt cgaagcatgg 2160
ccaggatgcc ggggtggagct ttccctcag ggtgatatgg agccgcaggc acagggccgc 2220
gactggctgc tggggctcaa caacctggac ggcctgcate gtgctctggg cctggcccat 2280
ggcgctctcg ctgaccttc gacgcgcgcg atacggctgg ctccgttgcg caatctaggt 2340
ctccgcgtcc tagtggtgga ggataacgcg atcaaccagt tgatcttgag ggaccagatg 2400
gaagcgctgg gctgcagcgt ggagctgctc ttcgatggtc gcgaggcgtt gctgcactgc 2460
cagacggcct gcttcgacgt ggtgctcacc gatatcaaca tgccgaacat gaacggatac 2520
gagctaaccg cggagctacg gcgcaaggg ttccggcagc cgatcatcgg cgcgacggcg 2580
aacgccatgc gtgaggagcg cgagcgctgc atgtccgcg ggatgaacga ttgcctggtc 2640
aaaccggtgg atctgaatgc ccttcagaac tgcttgatta atattctcaa ggtggatcga 2700
tga 2703

```

<210> 11

<211> 2664

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 11

```

atgggttttaa accatcaggt gtccctcagt cggcaagcta tgaatgtggc tatgtacgaa 60
ggcagccttt atttcgagca gcgcgaggcg ttgctcaatc acttgagcgg caatgtcgtg 120
cccttggccg cgggtagagc gctcgtcaac gaagcgccga acaatgtgag catcctgccg 180
ttgagtgcg gagggcgagg tctgctattg accgctcgca cgcctcggtga tctccgggaa 240
aagcggctgg cactgatgta tctggtcgat accgacaaag gccctctggt ttaccggctt 300
accgcgatg gtaggccctc ggcagcgata tccagcacga taaccaaaga ggtgtaccga 360
gccttgctgg cgaactcgtc ggcgccgtgt cactgggtga ctgacgggtg taccctcaa 420
cggctgtacc tttttgaatc cttaggcgat gagccgggcg aggggtgggt aggcctggag 480
attctcggcg aagacctcga ttcgatgttg cgccggaatg atgccgaaa ctacatgctg 540
ctggatcagc atgggcaggt cgtactcgct acggacgcag aggcgctggg gagcgggtgc 600
tcgcgagcgc ttttgctgg agacggcttc ggtttcatcg gtgctggccc actgccgcag 660
catatggtgc ttttccagca cgtggggtct tcgagctggg atctgatcta tcacatcgg 720
atcggctgcc tgttgctggc tctgtggctc cctctgttac ttgcctctgc gttggcactc 780
gcagtgcgca tccactgca ttggctgggt cggagcatcg agcgacgctt gatagagccc 840
gcaaagcgac gccttgaagc attgaaggag agcgaagcct tttcccggtc agttatccag 900
gccgcgcccc tcgogctgtg cgtgctgcgt cgtgccagc cgcagtggt cctggaaaat 960
ccccaggcgc cccaatggct gggatgagc gggcgattg cccacgacgc gccgagatgg 1020
atttccagg cgttcgcagg aggtgtgaag tgttctggag aagaactgga aaccgaggca 1080

```

```

gggctacatc ttcatotcaa ttacacgccc acccgctata acggtgaaga cgtattgttc 1140
tgccgcttca gtgaaatcag tgcacgcaag cggatggagg cggaaactggc tcgcgcaaaa 1200
tccctggcgg atgctgccaa tgaagccaag acgctgtttc tcgccaccat gagccatgaa 1260
atccgcacac ctctgtacgg catgcttggc acgcttgagc tgcttgggcg taccgagctg 1320
agtcggcagc agggccggtta cctaaaggca atccagcatt cctcgtcgac cctgctgcaa 1380
ctgatcagcg atgtgcttga cgtatccaag atagaggccg gccaactgga cctagagtgc 1440
gtggaattct ccccgctgga attgaccgaa gaggtcgtgc agtcgttcac cggtgccgcg 1500
caggccaagg ggctgcagtt gtatacctgc ctctctgagg agctgccgct gcgcatgcgg 1560
ggggccgagg cgtcgatccg gcagattctc aacaacctgc tgagcaacgc ggtgaagtgc 1620
accgacaatg gctatgtcaa cgtccacctg aaggccagcg ttgctgatgc cgaatgtgtg 1680
atgctgacct ggcagggtcaa cgataccggc atggggatca acgtcgagga tcagccgctg 1740
ctgttcgaac cgttctacca gatacggcgc tcgggcatc cggtcgcagg cacgggcctc 1800
ggcttgtcga tcagccagcg cctggcgag ctaatgaatg gcagtctgaa actggtcagt 1860
gagctggggg ttgggcagcag ctttagcctc aggtctccgc ttgagcggat cgcgatgcag 1920
gctgagccgc aggacctagc cgggtgcgcc gtccaagtgc tggcgccgtg ccgagacctc 1980
acggaatgcc tgtgtggctg gatctccgct tgggggtggaa gggccatggt cgcgacgctg 2040
aggctcgctg acgagtcgga cgcgacctgc ctgctggctg aagtgttact gctggagggg 2100
gcgcccagtg tcgaagcatg gccaggatgc cgggtggagc ttccctcca ggggtgatag 2160
gagccgcagg cacagggccg cgaactggct ctccggctca acaacctgga cggcctgcag 2220
cgtgctctgg gcttggccca tgggcgctc gctgatcctt cgacgcgccc gatacggctg 2280
gctccgcttg gcaatctagg tctccgctc ctatgggtgg aggataacgc gatcaaccag 2340
ttgatcttga gggaccagat ggaagcgctg ggctgcagcg ttgagctgct cttcgatggt 2400
cgcgaggcgt tgctgcaact ccagacggcc tgcttcgacg ttggtgctcac cgatatcaac 2460
atgccgaaca tgaacggata cgagctaacc gcggagctac ggcgccaagg gttccggcag 2520
ccgatcatcg gcgcgacggc gaacgcgatg cgtgaggagc gcgagcgctg catgtccgct 2580
gggatgaacg attgctggt caaacgggtg gatctgaatg cccttcagaa ctgcttgatt 2640
aatattctca aggtggatcg atga 2664

```

<210> 12

<211> 2625

<212> DNA

<213> Pseudomonas aeruginosa PA14

<400> 12

```

atgaatgtgg ctatgtacga agcgcagctt tatttcgagc agcgcgagggc gttgctcaat 60
cacttgagcg gcaatgtcgt gcccttggcc gcgggtagag cgctcgtcaa cgaagcgccg 120
aacaatgtga gcatcctgcc gttgagtgac ggaggcgag gtctgctatt gaccgctcgc 180
acgctcgggtg atctccggga aaagcggctg gcactgatgt atctggctga taccgacaaa 240
ggccctctgg tttaccggct taccgccgat ggtaggccct cggcagcgat atccagcacg 300
ataaccaaag aggtgtaccg agccttgctg gcgactccgt cggcgccctg tcaactgggtg 360
actgacgggtg gtacccctca acggctgtac ctttttgaat ccttaggcga tgagccgggc 420
gaggggtggc taggcctgga gattctcggc gaagacctcg attcgatgtt gcgccggaat 480
gatgccggaa actacatgct gctggatcag catgggcagg tcgtactcgc tacggacgca 540
gaggcgctgg ggagcgggtg gtgcgggacg cttttgcgtg gagacggctt cggtttcatc 600
ggtgctggcc cactgccgca gcatatgggt cttttccagc acgtggggtc ttcgagctgg 660
gatctgatct atcacatcgg tatcggtcgc ctggtgctgg ctctgtggct cctctgttta 720
cttgccctcg cgctggcact cgcagtcggc atcctactgc attggctggt gcggagcatc 780
gagcgacgct tgatagagcc cgcaaagcga cgccttgaag cattgaagga gagcgaagcc 840
ttttcccgct cagttatcca ggccgcgcc cgcgcgtgt gcgtgctgcg tcgtgccgac 900
gccgcagtggt tcctggaaaa tccccaggcg cgccaatggc tgggtgatag cgaggcgatt 960
gccacgagcg cgccgagatg gatttcccag gcgttcgcag gaggtgtgaa gtgttctgga 1020
gaagaactgg aaaccgaggg agggctacat cttcatctca attacacgcc caccgctat 1080
aacgggtgaag acgtattgtt ctgcgccttc agtgaaatca gtgcacgcaa gcggatggag 1140
gcggaactgg ctgcgcgcaa atccctggcg gatgctgcca atgaagccaa gacgctgttt 1200
ctgccacca tgagccatga aatccgcaca cctctgtacg gcatgcttgg cacgcttgag 1260
ctgcttggcg gtaccgagc caggccggtt acctaaaggc aatccagcat 1320
tcctcgtcga cctcgtgca actgatcagc gatgtgcttg acgtatccaa gatagaggcc 1380
ggccaactgg acctagagtg cgtggaattc tcccgcgtgg aattgaccga agaggctcgtg 1440

```

cagtcgttca	ccggtgcccg	gcaggccaag	gggctgcagt	tgtataacctg	cctctctgcg	1500
gagctgccgc	tgcgcatgcg	gggggcccgc	gcgtcgatcc	ggcagattct	caacaacctg	1560
ctgagcaacg	cggtgaagtt	caccgacaat	ggctatgtca	acgtccacct	gaaggccagc	1620
gtggtcgatg	ccgaatgtgt	gatgctgacc	tggcagggtca	acgataccgg	catggggatc	1680
aacgtcgagg	atcagccgcg	tctgttcgaa	ccgttctacc	agatacgccg	ctccgagcat	1740
ccggtcgcag	gcacgggcct	cggcttgctg	atcagccagc	gcctggcgca	gctaataaat	1800
ggcagctctga	aactgggtcag	tgagctgggg	ttgggcagca	gcttttagcct	caggcttccg	1860
cttgagcggg	tgcgcatgca	ggctgagccg	caggacctag	ccgggtgccc	cgtccaagtg	1920
ctgggcctcg	tccgcgacct	aacgggaatgc	ctgtgtggct	ggatctcccg	ctgggggtgga	1980
aggggcatgg	tgcgacgcgc	gaggtcgctg	gacgaggcgg	acgcgacctc	gctgctgggt	2040
gaagtgttac	tgcgggaggg	ggcgccgatg	ttcgaagcat	ggccaggatg	ccgggtggag	2100
ctttcccttc	agggtgatata	ggagccgcag	gcacaggggc	gcgactggct	gctcgggctc	2160
aacaacctgg	acggcctgca	tctgtctctg	ggcctggccc	atgggcgtct	cgtgatcct	2220
tgcagccgcg	cgatacggct	ggctccgttg	cgcaatctag	gtctccgcgt	cctagtgggtg	2280
gaggataaacg	cgatcaacca	gttgatcttg	agggaccaga	tggaaagcgt	gggctgcagc	2340
gtggagctgc	tcttcgatgg	tgcgagggcg	ttgctgcaact	gccagacggc	ctgcttcgac	2400
gtgggtgctca	ccgatataca	catgccgaac	atgaacggat	acgagctaac	cgcgagacta	2460
cgggcccaag	ggttccggca	gccgatcatc	ggcgcgacgg	cgaacgccat	gcgtgaggag	2520
cgcgagcgct	gcattgtccgc	cgggatgaac	gattgcctgg	tcaaaccggt	ggatctgaat	2580
gcccttcaga	actgcttgat	taattattctc	aaggtggatc	gatga		2625

<210> 13

<211> 2613

<212> DNA

<213> Pseudomonas aeruginosa PA14

<400> 13

atgtacgaag	cgcagcttta	tttcgagcag	cgcgaggcgt	tgtcfaatca	cttgagcggc	60
aatgtcgtgc	ccttggccgc	gggtagagcg	ctcgtcaacg	aagcgccgaa	caatgtgagc	120
atcctgccgt	tgagtgcagg	agggcgagggt	ctgctattga	ccgctcgcac	gctcggtgat	180
ctccgggaaa	agcggctggc	actgatgtat	ctggctcgata	ccgacaaagg	ccctctgggt	240
taccggctta	ccgcgcatgg	taggcctcgc	gcagcgatat	ccagcacgat	aaccaaagag	300
gtgtaccagag	ccttgctggc	gactccgtcg	cgccctgttc	actgggtgac	tgacggtggg	360
accctcaac	ggctgtacct	ttttgaatcc	ttaggcgatg	agccgggcca	gggggtggct	420
ggcctggaga	ttctcggcga	agacctcgat	tccatgttgc	gccggaatga	tgccggaaac	480
tacatgctgc	tggatcagca	tgggcagggtc	gtactcgcta	cggacgcaga	ggcgctgggg	540
agcgggtcgt	cgcggacgct	tttgctgga	gacggcttcg	gtttcatcgg	tgctggccca	600
ctgcgcgacg	atatgggtgct	tttccagcac	gtggggtctt	cgagctggga	tctgatctat	660
cacatcggtg	tccgtcgcct	gttgctggct	ctgtggctcc	ctctgttact	tgctctcgcg	720
ttggcactcg	cagtcggcat	cctactgcac	tggctgggtgc	ggagcatcga	gcgacgcttg	780
atagagcccc	caaagcgacg	ccttgaagca	ttgaaggaga	gcgaagcctt	ttcccgtagc	840
gttatccagg	ccgcgcccgt	cgcgctgtgc	gtgctgcgtc	gtgcgcgacg	cgcagtgggtc	900
ctggaaaatc	cccaggcgcg	ccaatggctg	gggtgatagcg	aggcgattgc	ccacgacgcg	960
ccgagatgga	tttccagggc	gttcgcaggga	gggtgtgaagt	gttctggaga	agaactggaa	1020
accgaggcag	ggctacatct	tcatctcaat	tacacgcccc	cccgtataaa	cggtgaagac	1080
gtattgttct	gcgccttcag	tgaatcaggt	gcacgcaagc	ggatggaggc	ggaactgggt	1140
cgcgcaaaat	ccctggcgga	tgctgccaat	gaagccaaga	cgctgtttct	cgccaccatg	1200
agccatgaaa	tccgcacacc	tctgtacggc	atgcttggca	cgtttgagct	gcttgggcgt	1260
accgagctga	gtcggcagca	ggccgggttac	ctaaaggcaa	tccagcatto	ctcgtcgacc	1320
ctgctgcaac	tgatcagcga	tgtgcttgac	gtatccaaga	tagaggccgg	ccaactggag	1380
ctagagctgcg	tgggaattctc	cccgctggaa	ttgaccgaag	aggctcgtgca	gtcgttcacc	1440
ggtagccgcg	aggccaagggt	gctgcagttg	tatacctgcc	tctctgcgga	gctgccgctg	1500
cgcattgcggg	gggcgcgggc	gtcgatccgg	cagattctca	acaacctgct	gagcaacgcg	1560
gtgaagtcca	ccgacaatgg	ctatgtcaac	gtccacctga	aggccagcgt	ggctcgatgcc	1620
gaatgtgtga	tgttgacctg	gcaggtcaac	ataccgggca	tggggatcaa	cgtcaggagg	1680
cagccgcgtc	gttctcgacc	gttctaccag	ataccgcgtc	ccgagcatcc	ggctcgaggc	1740
acgggcctcg	gcttgtcgat	cagccagcgc	ctggcgcgagc	taatgaatgg	cagctctgaaa	1800
ctgggtcagtg	agctgggggtt	gggcagcagc	tttagcctca	ggcttccgct	tgagcggatc	1860

```

gcgatgcagg ctgagccgca ggacctagcc ggggtgcgcc tccaagtgtc ggcgcctgtc 1920
cgcgacctaa cggaatgcct gtgtggctgg atctcccgct ggggtggaag ggccatggtc 1980
gcgacgccga ggtcgctgga cgaggcggac gcgacctcgc tgctggtcga agtggtactg 2040
ctggaggggg cgccgatgtt cgaagcatgg ccaggatgcc ggggtggagct ttccctcag 2100
gggtgatatgg agccgcaggc acaggggcgc gactggctgc tcgggctcaa caacctggac 2160
ggcctgcac gtgctctggg cctggcccat gggcgctctg ctgatccttc gacgccgccg 2220
atacggctgg ctccgttgcg caatctaggt ctccgcgtcc tagtggtgga ggataacgcg 2280
atcaaccagt tgatcttgag ggaccagatg gaagcgctgg gctgcagcgt ggagctgtc 2340
ttcgatggtc gcgaggcgtt gctgcactgc cagacggcct gcttcgacgt ggtgctcacc 2400
gatatcaaca tgccgaacat gaacggatac gagctaaccg cggagctacg gcgccaaggg 2460
ttccggcagc cgatcatcgg cgcgacggcg aacgccatgc gtgaggagcg cgagcgctgc 2520
atgtccgccg ggatgaacga ttgcctggtc aaaccgggtg atctgaatgc ccttcagaac 2580
tgcttgatta atattctcaa ggtggatcga tga 2613

```

<210> 14

<211> 2409

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 14

```

atgtatctgg tcgataccga caaaggccct ctggtttacc ggcttaccgc cgatggtagg 60
ccctcggcag cgatatccag cacgataacc aaagagggtg accgagcctt gctggcgact 120
ccgtcggcgc ctggttactg ggtgactgac ggtggtagcc ctcaacggct gtacctttt 180
gaatccttag gcgatgagcc gggcgagggg tggctaggcc tggagattct cggcgaagac 240
ctcgattcga tgttgccgg gaatgatgcc ggaaactaca tgctgctgga tcagcatggg 300
caggctgtac tcgctacgga cgcagaggcg ctggggagcg gtgcgtcgcg gacgcttttg 360
cgtggagacg gcttcggttt catcgggtgt ggcccaactgc cgcagcatat ggtgcttttc 420
cagcacgtgg ggtcttcgag ctgggatctg atctatcaca tcggtatcgg tcgctgttg 480
ctggctctgt ggctccctct gttacttgcc tctgcgttgg cactcgagct cggcatccta 540
ctgcattggc tgggtgaggag catcgagcga cgcttgatag agcccgcaa gcgacgcctt 600
gaagcattga aggagagcga agccttttcc cgtgcagtta tccaggccgc gcccgtcgcg 660
ctgtgcgtgc tgcgtcgtgc cgacgccgca gtggtcctgg aaaatcccca ggcgcgcca 720
tggctgggtg atagcgaggg gattgcccac gacgcgcga gatggatttc ccaggcgctt 780
gcaggaggtg tgaagtgttc tggagaagaa ctggaaaccg aggcagggct acatcttcat 840
ctcaattaca cgccaccgg ctataacggg gaagacgtat tgttctgcgc cttcagtga 900
atcagtgcac gcaagcggat ggaggcggaa ctggctcgcg caaaatccct ggcggatgct 960
gccaatgaag ccaagacgct gtttctcgcc accatgagcc atgaaatccg cacacctgt 1020
tacggcatgc ttggcacgct tgagctgctt gggcgtagcc agctgagtcg gcagcaggcc 1080
ggttacctaa aggcaatcca gcatctctcg tcgacctgc tgcaactgat cagcgatgtg 1140
cttgacgtat ccaagataga ggccggccaa ctggacctag agtgcggtga attctccccg 1200
ctggaattga ccgaagaggc cgtgcagtcg ttaccgggtg ccgcgcaggc caaggggctg 1260
cagttgtata cctgcctctc tgcggagctg ccgctgcgca tgcggggggc cgcgcgctcg 1320
atccggcaga ttctcaacaa cctgctgagc aacgcgggtg agttcaccga caatggctat 1380
gtcaacgtcc acctgaaggc cagcgtggtc gatgccgaat gtgtgatgct gacctggcag 1440
gtcaacgata ccggcatggg gatcaacgtc gaggatcagc cgcgtctgtt cgaaccgttc 1500
taccagatac gccgctccga gcatccggtc gcaggcacgg gcctcggctt gtcgatcagc 1560
cagcgctcgg cgcagctaat gaatggcagt ctgaaaactg tcagttagct ggggttgggc 1620
agcagcttta gcctcaggct tccgcttgag cggatcgcga tgcaggctga gccgcaggac 1680
ctagocgggt gcgccgtcca agtgctggcg cctgtccgcg acctaacgga atgcctgtgt 1740
ggctggatct cccgctgggg tggaggggcc atggtcgcga cgccgaggtc gctggacgag 1800
gcggacgcga cctcgctgct ggtcgaagtg ttactgctgg agggggcgcc gatgttcgaa 1860
gcatggccag gatgccgggt ggagctttcc cctcaggggtg atatggagcc gcaggcacag 1920
ggccgcgact ggctgctcgg gctcaacaac ctggacggcc tgcacgtgc tctgggctg 1980
gcccattggg gtctcgctga tccttcgacg ccgcgatac ggtcggctcc gttgcgcaat 2040
ctaggtctcc gcgtctagt ggtggaggat aacgcgatca accagttgat cttgagggac 2100
cagatggaag cgtgggctg cagcgtggag ctgctcttcg atggtcgcga ggcgtgtctg 2160
cactgccaga cggcctgctt cgacgtgggt ctcaccgata tcaacatgcc gaacatgaac 2220
ggatacgagc taaccgcgga gctacggcgc caaggggttc ggcagccgat catcgccgcg 2280

```

```

acggcgaaacg ccatgcgtga ggagcgcgag cgctgcatgt ccgccgggat gaacgattgc 2340
ctgggtcaaac cggtaggatct gaatgccctt cagaactgct tgattaatat tctcaagggtg 2400
gatcgatga                                     2409

```

<210> 15

<211> 2160

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 15

```

atgttgccgc ggaatgatgc cggaaactac atgtctgctgg atcagcatgg gcaggtcgta 60
ctcgctacgg acgcagaggg gctggggagc ggtgcgtcgc ggacgctttt gcgtggagac 120
ggcttcgggtt tcatcggtgc tggccactg ccgcagcata tgggtgctttt ccagcacgtg 180
gggtcttcga gctgggatct gatctatcac atcgggtatcg gtcgcctgtt gctggctctg 240
tggctccctc tgttacttgc ctctgcgttg gcaactcgag tcggcatcct actgcattgg 300
ctgggtgcgga gcacgcagcg acgcttgata gagcccgcaa agcgacgcct tgaagcattg 360
aaggagagcg aagccttttc ccgtgcagtt atccaggccg cgcctcgtcg gctgtgcgtg 420
ctgcgtcgtg ccgacgccgc agtggtcctg gaaaatcccc aggcgcgcca atggctgggt 480
gatagcgagg cgattgcca cgacgcgcgc agatggattt cccaggcggt cgcaggaggt 540
gtgaagtgtt ctggagaaga actggaaacc gaggcagggc tacatcttca tctcaattac 600
acgcccaccc gctataacgg tgaagacgta ttgttctgcg ccttcagtga aatcagtga 660
cgcaagcgga tggaggcgga actggctcgc gcaaaatccc tggcggatgc tgccaatgaa 720
gccaaagcgc tgtttctcgc caccatgagc catgaaatcc gcacacctct gtacggcatg 780
cttggcacgc ttgagctgct tgggcgtacc gagctgagtc ggcagcaggc cggttaccta 840
aaggcaatcc agcattcctc gtgcacctg ctgcaactga tcagcgatgt gcttgacgta 900
tccaagatag aggcgggcca actggacctg gactgcgtgg aattctcccc gctggaattg 960
accgaagagg tegtgcagtc gttcacgggt gccgcgcagg ccaaggggct gcagttgtat 1020
acctgcctct ctgcggagct gccgctgcgc atgcgggggg ccgcggcgct gatccggcag 1080
attctcaaca acctgctgag caacgcgggt aagttcaccc acaatggcta tgtcaacgta 1140
cacctgaagg ccagcgtggg cgatgccgaa tgtgtgatgc tgacctggca ggtcaacgat 1200
accggcatgg ggatcaacgt cgaggatcag ccgcgtctgt tcgaaccgtt ctaccagata 1260
cgccgctccg agcatccggg cgcaggcacg ggctcggct tgtcgatcag ccagcgcctg 1320
gcgcagctaa tgaatggcag tctgaaactg gtcagtgagc tgggggtggg cagcagcttt 1380
agcctcaggg ttccgcttga gcggatcgcg atgcaggctg agccgcagga cctagccggg 1440
tgccgcgtcc aagtgcctgg gcctgtccgc gacctaacgg aatgcctgtg tggctggatc 1500
tcccgtcggg gtggaagggg catggtcgcg acgccagggt cgctggacga ggcggacgcg 1560
acctcgctgc tggctgaagt gttactgctg gagggggcgc cgatgttcga agcatggcca 1620
ggatgccggg tggagctttc ccctcagggt gatatggagc cgcaggcaca gggccgcgac 1680
tggctcgctc ggctcaacaa cctggacggc ctgcactgtg ctctgggcct ggcccatggg 1740
cgtctcgctg atccttcgac gccgcgcgata cggctggctc cgttgcgcaa tctaggtctc 1800
cgcgctcctag tgggtggagga taacgcgata aaccagttga tcttgaggga ccagatggaa 1860
gcgctggggt gcagcgtgga gctgctcttc gatggtcgcg aggcgttgct gcactgccag 1920
acggcctgct tcgacgtggg gctcaccgat atcaacatgc cgaacatgaa cggatacgag 1980
ctaaccgcga agctacggcg ccaagggttc cggcagccga tcatcggcgc gacggcgaac 2040
gccatgcgtg aggagcgcga gcgctgcatg tccgccggga tgaacgattg cctgggtcaaa 2100
ccggtggatc tgaatgccct tcagaactgc ttgattaata ttctcaaggt ggatcgatga 2160

```

<210> 16

<211> 2130

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 16

```

atgtctgctgg atcagcatgg gcaggtcgta ctcgctacgg acgcagaggg gctggggagc 60
ggtgcgtcgc ggacgctttt gcgtggagac ggcttcgggt tcatcggtgc tggccactg 120
ccgcagcata tgggtctttt ccagcacgtg ggtgcttcga gctgggatct gatctatcac 180
atcgggtatcg gtcgcctgtt gctggctctg tggctccctc tgttacttgc ctctgcgttg 240

```

gcactcgcag	tccgcatcct	actgcattgg	ctggtgcgga	gcactcgcag	acgcttgata	300
gagcccgcaa	agcgacgcct	tgaagcattg	aaggagagcg	aagccttttc	ccgtgcagtt	360
atccaggccg	cgcccgctcg	gctgtgcgtg	ctgcgtcggt	ccgacgccc	agtggctctg	420
gaaaatcccc	aggcgcgcca	atggctgggt	gatagcgagg	cgattgccca	cgacgcgccc	480
agatggattt	cccaggcggt	cgcaggaggt	gtgaagtgtt	ctggagaaga	actggaaacc	540
gaggcagggc	tacatcttca	tctcaattac	acgcccaccc	gctataacgg	tgaagacgta	600
ttgttctgcg	ccttcagtg	aatcagtgca	cgcaagcgga	tggaggcgga	actggctcgc	660
gcaaaatccc	tggcgcatgc	tgccaatgaa	gccaaagcgc	tgtttctcgc	caccatgagc	720
catgaaatcc	gcacacctct	gtacggcatg	cttggcacgc	ttgagctgct	tgggctgacc	780
gagctgagtc	ggcagcaggc	cggttacct	aaggcaatcc	agcattcctc	gtcgaccctg	840
ctgcaactga	tcagcgatgt	gcttgacgta	tccaagatag	aggccggcca	actggacct	900
gagtgctgg	aattctcccc	gctggaattg	accgaagagg	tcgtgcagtc	gttcaccggg	960
gccgcgcagg	ccaaggggct	gcagttgtat	acctgcctct	ctgcggagct	gccgctgcgc	1020
atgccccggg	ccgcggcgct	gatccggcag	attctcaaca	acctgctgag	caacgcgggt	1080
aagttcaccc	acaatgggta	tgtcaacgtc	cacctgaagg	ccagcgtggg	cgatgccgaa	1140
tgtgtgatgc	tgacctggca	gggtcaacgat	accggcatgg	ggatcaacgt	cgaggatcag	1200
ccgctctgtg	tgaaccgtt	ctaccagata	cgccgctccg	agcatccggt	cgcaggcacg	1260
ggcctcggct	tgtcgatcag	ccagcgctcg	gcgcagctaa	tgaatggcag	tctgaaactg	1320
gtcagtgagc	tgggggttggg	cagcagcttt	agcctcaggc	ttccgcttga	gcggatcgcg	1380
atgcaggctg	agccgcagga	cctagccggg	tgcgccgtcc	aagtgcctgg	gcctgtccgc	1440
gacctaacgg	aatgcctgtg	tggctggatc	tcccgtggg	gtggaagggc	catggctcgc	1500
acgccgaggt	cgctggacga	ggcggacgcg	acctcgctgc	tggctgaagt	gttactgctg	1560
gagggggcgc	cgatgttcga	agcatggcca	ggatgcgggg	tggagctttc	ccctcagggt	1620
gatattggagc	cgcaggcaca	gggccgcgac	tggctgctcg	ggctcaacaa	cctggacggc	1680
ctgcatcgctg	ctctgggcct	ggcccatggg	cgctcgcctg	atccttcgac	gccgcccata	1740
cggctggctc	cgcttgccaa	tctaggtctc	cgcgtcctag	tggctggagg	taacgcgata	1800
aaccagttga	tcttgaggga	ccagatggaa	gcgctgggct	gcagcgtgga	gctgctcttc	1860
gatggctcgc	aggcgttgct	gcaactgccag	acggcctgct	tcgacgtggt	gctcaccgat	1920
atcaacatgc	cgaacatgaa	cggatacgag	ctaaccgcgg	agctacggcg	ccaaggggtc	1980
cggcagccga	tcacgcgcgc	gacggcgaa	gcatgcgtg	aggagcgcca	gcgctgcag	2040
tccgcgggga	tgaacgattg	cctgggtcaa	ccggtggatc	tgaatgccct	tcagaactgc	2100
ttgattaata	ttctcaagg	ggatcgatga				2130

<210> 17

<211> 2001

<212> DNA

<213> Pseudomonas aeruginosa PA14

<400> 17

atgggtgcttt	tccagcacgt	ggggtcttcg	agctggggtc	tgatctatca	catcggtatc	60
ggctgcctgt	tgttggtctc	gtggctccct	ctgttacttg	cctctgcgtt	ggcactcgca	120
gtcggcatcc	tactgcattg	gctggtgcgg	agcatcgagc	gacgcttgat	agagcccgcg	180
aagcgacgcc	ttgaagcatt	gaaggagagc	gaagcctttt	cccgtgcagt	tatccaggcc	240
gcgcccgtcg	cgctgtgcgt	gctgcgtcgt	gccgacgcgc	cagtggctct	ggaaaatccc	300
caggcgcgcc	aatggctggg	tgatagcgag	gcgattgccc	acgacgcgcc	gagatggatt	360
tcccaggcgt	tgcaggagg	tgtgaagtgt	tctggagaag	aactggaaac	cgaggcaggg	420
ctacatcttc	atctcaatta	cacgcccacc	cgctataacg	gtgaagacgt	attgttctgc	480
gccttcagtg	aaatcagtg	acgcaaagcg	atggaggcgg	aactggctcg	cgcaaatccc	540
ctggcgcatg	ctgccaatga	agccaagacg	ctgtttctcg	ccaccatgag	ccatgaaatc	600
cgcacacctc	tgtacggcat	gcttggcagc	cttgagctgc	ttgggctgac	cgagctgagt	660
cggcagcagg	ccggttacct	aaaggcaatc	cagcatctct	cgctcgacct	gctgcaactg	720
atcagcgatg	tgcttgacgt	atccaagata	gaggccggcc	aactggacct	agagtgcgtg	780
gaattctccc	cgctggaatt	gaccgaagag	gtcgtgcagt	cgttcacccg	tgcgcgcgag	840
gccaaggggc	tgcagttgta	tacctgcctc	tctgcggagc	tgcgcgtgcg	catgcggggg	900
gccgcggcgt	cgatccggca	gattctcaac	aacctgctga	gcaacgcggg	gaagttcacc	960
gacaatggct	atgtcaacgt	ccacctgaag	gccagcgtgg	tcgatgccga	atgtgtgatg	1020
ctgcactggc	agggtcaacg	taccggcatg	gggatacaac	tcgaggatca	gccgctctg	1080
ttcgaaccgt	tctaccagat	acgccgctcc	gagcatccgg	tcgcaggcac	gggcctcggc	1140

```

ttgtcgatca gccagcgctt ggccgagcta atgaatggca gtctgaaact ggtcagttag 1200
ctgggggttg gcagcagctt tagcctcagg ctcccgcttg agcggatcgc gatgcaggct 1260
gagccgcagg acctagccgg gtgcccgttc caagtgtctg ccctgtccg cgacctaacg 1320
gaatgcctgt gtggctggat ctcccgtctg ggtggaaggg ccatggtcgc gacggcagg 1380
tcgctggacg agggcgacgc gacctcgctg ctggtcgaag tgttactgct ggagggggcg 1440
ccgatgttcg aagcatggcc aggatgccgg gtggagcttt cccctcaggg tgatatggag 1500
ccgcaggcac agggccgcga ctggctgctc gggctcaaca acctggacgg cctgcatcgt 1560
gctctggggc tggcccatgg gcgtctcgct gatccttcga cgccgccgat acggctggct 1620
ccgttgccga atctaggtct ccgcgtccta gtggtggagg ataacgcgat caaccagttg 1680
atcttgaggg accagatgga agcgtggggc tgcagcgtgg agctgctctt cgatggctcg 1740
gaggcgcttg tgcactgcca gacggcctgc ttcgacgtgg tgctcaccga tatcaacatg 1800
ccgaacatga acggatacga gctaaccgcg gagctacggc gccaaaggggt ccggcagccg 1860
atcatcgggc cgacggcgaa cgccatgcgt gaggagcgcg agcgtgcat gtccgcccgg 1920
atgaacgatt gcctggtcaa accggtggat ctgaatgccc ttcagaactg cttgattaat 1980
attctcaagg tggatcgatg a 2001

```

<210> 18

<211> 1491

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 18

```

atggaggcgg aactggctcg cgcaaaatcc ctggcgggat ctgccaatga agccaagacg 60
ctgtttctcg ccaccatgag ccatgaaatc cgcacacctc tgtacggcat gcttggcacg 120
cttgagctgc ttgggcgtac cgagctgagt cggcagcagg ccggttacct aaaggcaatc 180
cagcatctct cgtcgaccct gctgcaactg atcagcgatg tgcttgacgt atccaagata 240
gaggccggcc aactggacct agagtgcgtg gaattctccc cgttggaatt gaccgaagag 300
gtcgtgcagt cgttcaccgg tgcgcgcgag gccaaagggg tgcagttaga tacctgcctc 360
tctgoggagc tgcgcgtgcg catgccccgg gccgcggcgt cgatccggca gattctcaac 420
aacctgtcga gcaacgcggt gaagttcacc gacaatggct atgtcaacgt ccacctgaag 480
gccagcgtgg tcatgcccga atgtgtgatg ctgacctggc aggtcaacga taccggcatg 540
gggatcaacg tccaggatca gccgcgtctg ttcgaacctg tctaccagat acgccgctcc 600
gagcatccgg tgcagggcac gggcctcggc ttgtcgatca gccagcgctt ggccgagcta 660
atgaatggca gtctgaaact ggtcagttag ctgggggttg gcagcagctt tagcctcagg 720
cttccgcttg agcggatcgc gatgcaggct gagccgcagg acctagccgg gtgcccgttc 780
caagtgtctg ccctgtccg cgacctaacg gaatgcctgt gtggctggat ctcccgtctg 840
ggtggaaggg ccatggtcgc gacgccgagg tccgtggacg agggcgacgc gacctcgctg 900
ctggtcgaag tgttactgct ggagggggcg ccgatgttcg aagcatggcc aggatggcgg 960
gtggagcttt cccctcaggg cctgcatcgt gctctggggc tggcccatgg gcgtctcgct 1020
gggctcaaca acctggacgg cctgcatcgt gctctggggc tggcccatgg gcgtctcgct 1080
gatccttcga cgccgccgat acggctggct ccgttgcgca atctaggtct ccgcgtccta 1140
gtggtggagg ataacgcgat caaccagttg atcttgaggg accagatgga agcgtggggc 1200
tgcagcgtgg agctgctctt cgatggctgc gagcggttgc tgcactgcca gacggcctgc 1260
ttcgacgtgg tgctcaccga tatcaacatg ccgaacatga acggatacga gctaaccgcg 1320
gagctacggc gccaaaggggt ccggcagccg atcatcgggc cgacggcgaa cgccatgcgt 1380
gaggagcgcg agcgtgcat gtccgcccgg atgaacgatt gcctggtcaa accggtggat 1440
ctgaatgccc ttcagaactg cttgattaat attctcaagg tggatcgatg a 1491

```

<210> 19

<211> 931

<212> PRT

<213> *Pseudomonas aeruginosa* PA14

<400> 19

```

Met Lys Leu Lys Asn Phe Leu Gln Pro Phe Asp Ser Gly Phe Ser Thr
1           5           10           15
Pro Ser Ala Ala Leu Lys Leu Leu Arg Met Leu Gly Gly Ala Leu Met
20           25           30

```

Leu Cys Val Leu Cys Ser Leu Ile Phe Ser Val Ser Met Val Leu Asn
 35 40 45
 His Gln Val Ser Leu Ser Arg Gln Ala Met Asn Val Ala Met Tyr Glu
 50 55 60
 Ala Gln Leu Tyr Phe Glu Gln Arg Glu Ala Leu Leu Asn His Leu Ser
 65 70 75 80
 Gly Asn Val Val Pro Leu Ala Ala Gly Arg Ala Leu Val Asn Glu Ala
 85 90 95
 Pro Asn Asn Val Ser Ile Leu Pro Leu Ser Asp Gly Gly Arg Gly Leu
 100 105 110
 Leu Leu Thr Ala Arg Thr Leu Gly Asp Leu Arg Glu Lys Arg Leu Ala
 115 120 125
 Leu Met Tyr Leu Val Asp Thr Asp Lys Gly Pro Leu Val Tyr Arg Leu
 130 135 140
 Thr Ala Asp Gly Arg Pro Ser Ala Ala Ile Ser Ser Thr Ile Thr Lys
 145 150 155 160
 Glu Val Tyr Arg Ala Leu Leu Ala Thr Pro Ser Ala Pro Val His Trp
 165 170 175
 Val Thr Asp Gly Gly Thr Pro Gln Arg Leu Tyr Leu Phe Glu Ser Leu
 180 185 190
 Gly Asp Glu Pro Gly Glu Gly Trp Leu Gly Leu Glu Ile Leu Gly Glu
 195 200 205
 Asp Leu Asp Ser Met Leu Arg Arg Asn Asp Ala Gly Asn Tyr Met Leu
 210 215 220
 Leu Asp Gln His Gly Gln Val Val Leu Ala Thr Asp Ala Glu Ala Leu
 225 230 235 240
 Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp Gly Phe Gly Phe
 245 250 255
 Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu Phe Gln His Val
 260 265 270
 Gly Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly Ile Gly Arg Leu
 275 280 285
 Leu Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser Ala Leu Ala Leu
 290 295 300
 Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser Ile Glu Arg Arg
 305 310 315 320
 Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu Lys Glu Ser Glu
 325 330 335
 Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val Ala Leu Cys Val
 340 345 350
 Leu Arg Arg Ala Asp Ala Ala Val Val Leu Glu Asn Pro Gln Ala Arg
 355 360 365
 Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp Ala Pro Arg Trp
 370 375 380
 Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser Gly Glu Glu Leu
 385 390 395 400
 Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr Thr Pro Thr Arg
 405 410 415
 Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser Glu Ile Ser Ala
 420 425 430
 Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu Ala Asp
 435 440 445
 Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr Met Ser His Glu
 450 455 460
 Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu Leu Gly
 465 470 475 480
 Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala Ile Gln
 485 490 495

His	Ser	Ser	Ser	Thr	Leu	Leu	Gln	Leu	Ile	Ser	Asp	Val	Leu	Asp	Val
			500					505					510		
Ser	Lys	Ile	Glu	Ala	Gly	Gln	Leu	Asp	Leu	Glu	Cys	Val	Glu	Phe	Ser
	515						520					525			
Pro	Leu	Glu	Leu	Thr	Glu	Glu	Val	Val	Gln	Ser	Phe	Thr	Gly	Ala	Ala
	530					535					540				
Gln	Ala	Lys	Gly	Leu	Gln	Leu	Tyr	Thr	Cys	Leu	Ser	Ala	Glu	Leu	Pro
545					550					555					560
Leu	Arg	Met	Arg	Gly	Ala	Ala	Ala	Ser	Ile	Arg	Gln	Ile	Leu	Asn	Asn
				565					570					575	
Leu	Leu	Ser	Asn	Ala	Val	Lys	Phe	Thr	Asp	Asn	Gly	Tyr	Val	Asn	Val
			580						585				590		
His	Leu	Lys	Ala	Ser	Val	Val	Asp	Ala	Glu	Cys	Val	Met	Leu	Thr	Trp
	595						600					605			
Gln	Val	Asn	Asp	Thr	Gly	Met	Gly	Ile	Asn	Val	Glu	Asp	Gln	Pro	Arg
	610					615					620				
Leu	Phe	Glu	Pro	Phe	Tyr	Gln	Ile	Arg	Arg	Ser	Glu	His	Pro	Val	Ala
625					630					635					640
Gly	Thr	Gly	Leu	Gly	Leu	Ser	Ile	Ser	Gln	Arg	Leu	Ala	Gln	Leu	Met
				645					650					655	
Asn	Gly	Ser	Leu	Lys	Leu	Val	Ser	Glu	Leu	Gly	Leu	Gly	Ser	Ser	Phe
			660					665					670		
Ser	Leu	Arg	Leu	Pro	Leu	Glu	Arg	Ile	Ala	Met	Gln	Ala	Glu	Pro	Gln
	675						680					685			
Asp	Leu	Ala	Gly	Cys	Ala	Val	Gln	Val	Leu	Ala	Pro	Val	Arg	Asp	Leu
	690					695					700				
Thr	Glu	Cys	Leu	Cys	Gly	Trp	Ile	Ser	Arg	Trp	Gly	Gly	Arg	Ala	Met
705					710					715					720
Val	Ala	Thr	Pro	Arg	Ser	Leu	Asp	Glu	Ala	Asp	Ala	Thr	Ser	Leu	Leu
				725					730					735	
Val	Glu	Val	Leu	Leu	Leu	Glu	Gly	Ala	Pro	Met	Phe	Glu	Ala	Trp	Pro
			740					745					750		
Gly	Cys	Arg	Val	Glu	Leu	Ser	Pro	Gln	Gly	Asp	Met	Glu	Pro	Gln	Ala
	755						760					765			
Gln	Gly	Arg	Asp	Trp	Leu	Leu	Gly	Leu	Asn	Asn	Leu	Asp	Gly	Leu	His
	770					775					780				
Arg	Ala	Leu	Gly	Leu	Ala	His	Gly	Arg	Leu	Ala	Asp	Pro	Ser	Thr	Pro
785					790					795					800
Pro	Ile	Arg	Leu	Ala	Pro	Leu	Arg	Asn	Leu	Gly	Leu	Arg	Val	Leu	Val
				805					810					815	
Val	Glu	Asp	Asn	Ala	Ile	Asn	Gln	Leu	Ile	Leu	Arg	Asp	Gln	Met	Glu
			820					825					830		
Ala	Leu	Gly	Cys	Ser	Val	Glu	Leu	Leu	Phe	Asp	Gly	Arg	Glu	Ala	Leu
	835						840					845			
Leu	His	Cys	Gln	Thr	Ala	Cys	Phe	Asp	Val	Val	Leu	Thr	Asp	Ile	Asn
	850					855					860				
Met	Pro	Asn	Met	Asn	Gly	Tyr	Glu	Leu	Thr	Ala	Glu	Leu	Arg	Arg	Gln
865					870					875					880
Gly	Phe	Arg	Gln	Pro	Ile	Ile	Gly	Ala	Thr	Ala	Asn	Ala	Met	Arg	Glu
				885					890					895	
Glu	Arg	Glu	Arg	Cys	Met	Ser	Ala	Gly	Met	Asn	Asp	Cys	Leu	Val	Lys
			900					905					910		
Pro	Val	Asp	Leu	Asn	Ala	Leu	Gln	Asn	Cys	Leu	Ile	Asn	Ile	Leu	Lys
		915					920					925			
Val	Asp	Arg													
															930

<210> 20
 <211> 906
 <212> PRT
 <213> Pseudomonas aeruginosa PA14

<400> 20
 Met Leu Gly Gly Ala Leu Met Leu Cys Val Leu Cys Ser Leu Ile Phe
 1 5 10 15
 Ser Val Ser Met Val Leu Asn His Gln Val Ser Leu Ser Arg Gln Ala
 20 25 30
 Met Asn Val Ala Met Tyr Glu Ala Gln Leu Tyr Phe Glu Gln Arg Glu
 35 40 45
 Ala Leu Leu Asn His Leu Ser Gly Asn Val Val Pro Leu Ala Ala Gly
 50 55 60
 Arg Ala Leu Val Asn Glu Ala Pro Asn Asn Val Ser Ile Leu Pro Leu
 65 70 75 80
 Ser Asp Gly Gly Arg Gly Leu Leu Leu Thr Ala Arg Thr Leu Gly Asp
 85 90 95
 Leu Arg Glu Lys Arg Leu Ala Leu Met Tyr Leu Val Asp Thr Asp Lys
 100 105 110
 Gly Pro Leu Val Tyr Arg Leu Thr Ala Asp Gly Arg Pro Ser Ala Ala
 115 120 125
 Ile Ser Ser Thr Ile Thr Lys Glu Val Tyr Arg Ala Leu Leu Ala Thr
 130 135 140
 Pro Ser Ala Pro Val His Trp Val Thr Asp Gly Gly Thr Pro Gln Arg
 145 150 155 160
 Leu Tyr Leu Phe Glu Ser Leu Gly Asp Glu Pro Gly Glu Gly Trp Leu
 165 170 175
 Gly Leu Glu Ile Leu Gly Glu Asp Leu Asp Ser Met Leu Arg Arg Asn
 180 185 190
 Asp Ala Gly Asn Tyr Met Leu Leu Asp Gln His Gly Gln Val Val Leu
 195 200 205
 Ala Thr Asp Ala Glu Ala Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu
 210 215 220
 Arg Gly Asp Gly Phe Gly Phe Ile Gly Ala Gly Pro Leu Pro Gln His
 225 230 235 240
 Met Val Leu Phe Gln His Val Gly Ser Ser Trp Asp Leu Ile Tyr
 245 250 255
 His Ile Gly Ile Gly Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu
 260 265 270
 Leu Ala Ser Ala Leu Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu
 275 280 285
 Val Arg Ser Ile Glu Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu
 290 295 300
 Glu Ala Leu Lys Glu Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala
 305 310 315 320
 Ala Pro Val Ala Leu Cys Val Leu Arg Arg Ala Asp Ala Val Val
 325 330 335
 Leu Glu Asn Pro Gln Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile
 340 345 350
 Ala His Asp Ala Pro Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val
 355 360 365
 Lys Cys Ser Gly Glu Glu Leu Glu Thr Glu Ala Gly Leu His Leu His
 370 375 380
 Leu Asn Tyr Thr Pro Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys
 385 390 395 400
 Ala Phe Ser Glu Ile Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala
 405 410 415

Arg	Ala	Lys	Ser	Leu	Ala	Asp	Ala	Ala	Asn	Glu	Ala	Lys	Thr	Leu	Phe
			420					425					430		
Leu	Ala	Thr	Met	Ser	His	Glu	Ile	Arg	Thr	Pro	Leu	Tyr	Gly	Met	Leu
		435					440					445			
Gly	Thr	Leu	Glu	Leu	Leu	Gly	Arg	Thr	Glu	Leu	Ser	Arg	Gln	Gln	Ala
		450				455					460				
Gly	Tyr	Leu	Lys	Ala	Ile	Gln	His	Ser	Ser	Ser	Thr	Leu	Leu	Gln	Leu
465					470					475					480
Ile	Ser	Asp	Val	Leu	Asp	Val	Ser	Lys	Ile	Glu	Ala	Gly	Gln	Leu	Asp
			485						490					495	
Leu	Glu	Cys	Val	Glu	Phe	Ser	Pro	Leu	Glu	Leu	Thr	Glu	Glu	Val	Val
			500					505					510		
Gln	Ser	Phe	Thr	Gly	Ala	Ala	Gln	Ala	Lys	Gly	Leu	Gln	Leu	Tyr	Thr
		515					520					525			
Cys	Leu	Ser	Ala	Glu	Leu	Pro	Leu	Arg	Met	Arg	Gly	Ala	Ala	Ala	Ser
		530				535					540				
Ile	Arg	Gln	Ile	Leu	Asn	Asn	Leu	Leu	Ser	Asn	Ala	Val	Lys	Phe	Thr
545					550					555					560
Asp	Asn	Gly	Tyr	Val	Asn	Val	His	Leu	Lys	Ala	Ser	Val	Val	Asp	Ala
			565						570					575	
Glu	Cys	Val	Met	Leu	Thr	Trp	Gln	Val	Asn	Asp	Thr	Gly	Met	Gly	Ile
		580					585						590		
Asn	Val	Glu	Asp	Gln	Pro	Arg	Leu	Phe	Glu	Pro	Phe	Tyr	Gln	Ile	Arg
		595					600					605			
Arg	Ser	Glu	His	Pro	Val	Ala	Gly	Thr	Gly	Leu	Gly	Leu	Ser	Ile	Ser
		610				615					620				
Gln	Arg	Leu	Ala	Gln	Leu	Met	Asn	Gly	Ser	Leu	Lys	Leu	Val	Ser	Glu
625					630					635					640
Leu	Gly	Leu	Gly	Ser	Ser	Phe	Ser	Leu	Arg	Leu	Pro	Leu	Glu	Arg	Ile
			645						650					655	
Ala	Met	Gln	Ala	Glu	Pro	Gln	Asp	Leu	Ala	Gly	Cys	Ala	Val	Gln	Val
		660						665					670		
Leu	Ala	Pro	Val	Arg	Asp	Leu	Thr	Glu	Cys	Leu	Cys	Gly	Trp	Ile	Ser
		675				680						685			
Arg	Trp	Gly	Gly	Arg	Ala	Met	Val	Ala	Thr	Pro	Arg	Ser	Leu	Asp	Glu
		690				695					700				
Ala	Asp	Ala	Thr	Ser	Leu	Val	Glu	Val	Leu	Leu	Leu	Glu	Gly	Ala	
705					710				715						720
Pro	Met	Phe	Glu	Ala	Trp	Pro	Gly	Cys	Arg	Val	Glu	Leu	Ser	Pro	Gln
			725						730					735	
Gly	Asp	Met	Glu	Pro	Gln	Ala	Gln	Gly	Arg	Asp	Trp	Leu	Leu	Gly	Leu
		740						745					750		
Asn	Asn	Leu	Asp	Gly	Leu	His	Arg	Ala	Leu	Gly	Leu	Ala	His	Gly	Arg
		755					760					765			
Leu	Ala	Asp	Pro	Ser	Thr	Pro	Pro	Ile	Arg	Leu	Ala	Pro	Leu	Arg	Asn
		770				775					780				
Leu	Gly	Leu	Arg	Val	Leu	Val	Val	Glu	Asp	Asn	Ala	Ile	Asn	Gln	Leu
785					790					795					800
Ile	Leu	Arg	Asp	Gln	Met	Glu	Ala	Leu	Gly	Cys	Ser	Val	Glu	Leu	Leu
			805						810					815	
Phe	Asp	Gly	Arg	Glu	Ala	Leu	Leu	His	Cys	Gln	Thr	Ala	Cys	Phe	Asp
			820					825					830		
Val	Val	Leu	Thr	Asp	Ile	Asn	Met	Pro	Asn	Met	Asn	Gly	Tyr	Glu	Leu
		835				840						845			
Thr	Ala	Glu	Leu	Arg	Arg	Gln	Gly	Phe	Arg	Gln	Pro	Ile	Ile	Gly	Ala
		850				855					860				
Thr	Ala	Asn	Ala	Met	Arg	Glu	Glu	Arg	Glu	Arg	Cys	Met	Ser	Ala	Gly
865					870					875					880

Met Asn Asp Cys Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn
 885 890 895
 Cys Leu Ile Asn Ile Leu Lys Val Asp Arg
 900 905

<210> 21
 <211> 900
 <212> PRT
 <213> Pseudomonas aeruginosa PA14

<400> 21
 Met Leu Cys Val Leu Cys Ser Leu Ile Phe Ser Val Ser Met Val Leu
 1 5 10 15
 Asn His Gln Val Ser Leu Ser Arg Gln Ala Met Asn Val Ala Met Tyr
 20 25 30
 Glu Ala Gln Leu Tyr Phe Glu Gln Arg Glu Ala Leu Leu Asn His Leu
 35 40 45
 Ser Gly Asn Val Val Pro Leu Ala Ala Gly Arg Ala Leu Val Asn Glu
 50 55 60
 Ala Pro Asn Asn Val Ser Ile Leu Pro Leu Ser Asp Gly Gly Arg Gly
 65 70 75 80
 Leu Leu Leu Thr Ala Arg Thr Leu Gly Asp Leu Arg Glu Lys Arg Leu
 85 90 95
 Ala Leu Met Tyr Leu Val Asp Thr Asp Lys Gly Pro Leu Val Tyr Arg
 100 105 110
 Leu Thr Ala Asp Gly Arg Pro Ser Ala Ala Ile Ser Ser Thr Ile Thr
 115 120 125
 Lys Glu Val Tyr Arg Ala Leu Leu Ala Thr Pro Ser Ala Pro Val His
 130 135 140
 Trp Val Thr Asp Gly Gly Thr Pro Gln Arg Leu Tyr Leu Phe Glu Ser
 145 150 155 160
 Leu Gly Asp Glu Pro Gly Glu Gly Trp Leu Gly Leu Glu Ile Leu Gly
 165 170 175
 Glu Asp Leu Asp Ser Met Leu Arg Arg Asn Asp Ala Gly Asn Tyr Met
 180 185 190
 Leu Leu Asp Gln His Gly Gln Val Val Leu Ala Thr Asp Ala Glu Ala
 195 200 205
 Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp Gly Phe Gly
 210 215 220
 Phe Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu Phe Gln His
 225 230 235 240
 Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly Ile Gly Arg
 245 250 255
 Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser Ala Leu Ala
 260 265 270
 Leu Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser Ile Glu Arg
 275 280 285
 Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu Lys Glu Ser
 290 295 300
 Glu Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val Ala Leu Cys
 305 310 315 320
 Val Leu Arg Arg Ala Asp Ala Ala Val Val Leu Glu Asn Pro Gln Ala
 325 330 335
 Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp Ala Pro Arg
 340 345 350
 Trp Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser Gly Glu Glu
 355 360 365

```

Leu Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr Thr Pro Thr
  370          375          380
Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser Glu Ile Ser
385          390          395          400
Ala Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu Ala
          405          410          415
Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr Met Ser His
          420          425          430
Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu Leu
          435          440          445
Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala Ile
          450          455          460
Gln His Ser Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp Val Leu Asp
465          470          475          480
Val Ser Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys Val Glu Phe
          485          490          495
Ser Pro Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe Thr Gly Ala
          500          505          510
Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser Ala Glu Leu
          515          520          525
Pro Leu Arg Met Arg Gly Ala Ala Ala Ser Ile Arg Gln Ile Leu Asn
530          535          540
Asn Leu Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly Tyr Val Asn
545          550          555          560
Val His Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val Met Leu Thr
          565          570          575
Trp Gln Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu Asp Gln Pro
          580          585          590
Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu His Pro Val
          595          600          605
Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln Leu
          610          615          620
Met Asn Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu Gly Ser Ser
625          630          635          640
Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln Ala Glu Pro
          645          650          655
Gln Asp Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro Val Arg Asp
          660          665          670
Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg Ala
          675          680          685
Met Val Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser Leu
          690          695          700
Leu Val Glu Val Leu Leu Leu Glu Gly Ala Pro Met Phe Glu Ala Trp
705          710          715          720
Pro Gly Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met Glu Pro Gln
          725          730          735
Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu Asn Asn Leu Asp Gly Leu
          740          745          750
His Arg Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp Pro Ser Thr
          755          760          765
Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu Arg Val Leu
          770          775          780
Val Val Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg Asp Gln Met
785          790          795          800
Glu Ala Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly Arg Glu Ala
          805          810          815
Leu Leu His Cys Gln Thr Ala Cys Phe Asp Val Val Leu Thr Asp Ile
          820          825          830

```

Asn Met Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu Leu Arg Arg
 835 840 845
 Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn Ala Met Arg
 850 855 860
 Glu Glu Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp Cys Leu Val
 865 870 875 880
 Lys Pro Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile Asn Ile Leu
 885 890 895
 Lys Val Asp Arg
 900

<210> 22
 <211> 887
 <212> PRT
 <213> Pseudomonas aeruginosa PA14

<400> 22
 Met Val Leu Asn His Gln Val Ser Leu Ser Arg Gln Ala Met Asn Val
 1 5 10 15
 Ala Met Tyr Glu Ala Gln Leu Tyr Phe Glu Gln Arg Glu Ala Leu Leu
 20 25 30
 Asn His Leu Ser Gly Asn Val Val Pro Leu Ala Ala Gly Arg Ala Leu
 35 40 45
 Val Asn Glu Ala Pro Asn Asn Val Ser Ile Leu Pro Leu Ser Asp Gly
 50 55 60
 Gly Arg Gly Leu Leu Leu Thr Ala Arg Thr Leu Gly Asp Leu Arg Glu
 65 70 75 80
 Lys Arg Leu Ala Leu Met Tyr Leu Val Asp Thr Asp Lys Gly Pro Leu
 85 90 95
 Val Tyr Arg Leu Thr Ala Asp Gly Arg Pro Ser Ala Ala Ile Ser Ser
 100 105 110
 Thr Ile Thr Lys Glu Val Tyr Arg Ala Leu Leu Ala Thr Pro Ser Ala
 115 120 125
 Pro Val His Trp Val Thr Asp Gly Gly Thr Pro Gln Arg Leu Tyr Leu
 130 135 140
 Phe Glu Ser Leu Gly Asp Glu Pro Gly Glu Gly Trp Leu Gly Leu Glu
 145 150 155 160
 Ile Leu Gly Glu Asp Leu Asp Ser Met Leu Arg Arg Asn Asp Ala Gly
 165 170 175
 Asn Tyr Met Leu Leu Asp Gln His Gly Gln Val Val Leu Ala Thr Asp
 180 185 190
 Ala Glu Ala Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp
 195 200 205
 Gly Phe Gly Phe Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu
 210 215 220
 Phe Gln His Val Gly Ser Ser Trp Asp Leu Ile Tyr His Ile Gly
 225 230 235 240
 Ile Gly Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser
 245 250 255
 Ala Leu Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser
 260 265 270
 Ile Glu Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu
 275 280 285
 Lys Glu Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val
 290 295 300
 Ala Leu Cys Val Leu Arg Arg Ala Asp Ala Val Val Leu Glu Asn
 305 310 315 320

Pro Gln Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp
 325 330 335
 Ala Pro Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser
 340 345 350
 Gly Glu Glu Leu Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr
 355 360 365
 Thr Pro Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser
 370 375 380
 Glu Ile Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys
 385 390 395 400
 Ser Leu Ala Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr
 405 410 415
 Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu
 420 425 430
 Glu Leu Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu
 435 440 445
 Lys Ala Ile Gln His Ser Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp
 450 455 460
 Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys
 465 470 475 480
 Val Glu Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe
 485 490 495
 Thr Gly Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser
 500 505 510
 Ala Glu Leu Pro Leu Arg Met Arg Gly Ala Ala Ala Ser Ile Arg Gln
 515 520 525
 Ile Leu Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly
 530 535 540
 Tyr Val Asn Val His Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val
 545 550 555 560
 Met Leu Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu
 565 570 575
 Asp Gln Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu
 580 585 590
 His Pro Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu
 595 600 605
 Ala Gln Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu
 610 615 620
 Gly Ser Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln
 625 630 635 640
 Ala Glu Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro
 645 650 655
 Val Arg Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly
 660 665 670
 Gly Arg Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala
 675 680 685
 Thr Ser Leu Leu Val Glu Val Leu Leu Glu Gly Ala Pro Met Phe
 690 695 700
 Glu Ala Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met
 705 710 715 720
 Glu Pro Gln Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu Asn Asn Leu
 725 730 735
 Asp Gly Leu His Arg Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp
 740 745 750
 Pro Ser Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu
 755 760 765
 Arg Val Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg
 770 775 780

```

Asp Gln Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly
785              790              795              800
Arg Glu Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp Val Val Leu
              805              810              815
Thr Asp Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu
              820              825              830
Leu Arg Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn
              835              840              845
Ala Met Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp
              850              855              860
Cys Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile
865              870              875              880
Asn Ile Leu Lys Val Asp Arg
              885

```

<210> 23

<211> 874

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 23

```

Met Asn Val Ala Met Tyr Glu Ala Gln Leu Tyr Phe Glu Gln Arg Glu
1              5              10              15
Ala Leu Leu Asn His Leu Ser Gly Asn Val Val Pro Leu Ala Ala Gly
              20              25              30
Arg Ala Leu Val Asn Glu Ala Pro Asn Asn Val Ser Ile Leu Pro Leu
              35              40              45
Ser Asp Gly Gly Arg Gly Leu Leu Leu Thr Ala Arg Thr Leu Gly Asp
              50              55              60
Leu Arg Glu Lys Arg Leu Ala Leu Met Tyr Leu Val Asp Thr Asp Lys
65              70              75              80
Gly Pro Leu Val Tyr Arg Leu Thr Ala Asp Gly Arg Pro Ser Ala Ala
              85              90              95
Ile Ser Ser Thr Ile Thr Lys Glu Val Tyr Arg Ala Leu Leu Ala Thr
              100              105              110
Pro Ser Ala Pro Val His Trp Val Thr Asp Gly Gly Thr Pro Gln Arg
              115              120              125
Leu Tyr Leu Phe Glu Ser Leu Gly Asp Glu Pro Gly Glu Gly Trp Leu
              130              135              140
Gly Leu Glu Ile Leu Gly Glu Asp Leu Asp Ser Met Leu Arg Arg Asn
145              150              155              160
Asp Ala Gly Asn Tyr Met Leu Leu Asp Gln His Gly Gln Val Val Leu
              165              170              175
Ala Thr Asp Ala Glu Ala Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu
              180              185              190
Arg Gly Asp Gly Phe Gly Phe Ile Gly Ala Gly Pro Leu Pro Gln His
              195              200              205
Met Val Leu Phe Gln His Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr
210              215              220
His Ile Gly Ile Gly Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu
225              230              235              240
Leu Ala Ser Ala Leu Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu
              245              250              255
Val Arg Ser Ile Glu Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu
              260              265              270
Glu Ala Leu Lys Glu Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala
275              280              285

```



```

Ala Pro Val Ala Leu Cys Val Leu Arg Arg Ala Asp Ala Ala Val Val
 290          295          300
Leu Glu Asn Pro Gln Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile
305          310          315          320
Ala His Asp Ala Pro Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val
          325          330          335
Lys Cys Ser Gly Glu Glu Leu Glu Thr Glu Ala Gly Leu His Leu His
          340          345          350
Leu Asn Tyr Thr Pro Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys
          355          360          365
Ala Phe Ser Glu Ile Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala
          370          375          380
Arg Ala Lys Ser Leu Ala Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe
385          390          395          400
Leu Ala Thr Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu
          405          410          415
Gly Thr Leu Glu Leu Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala
          420          425          430
Gly Tyr Leu Lys Ala Ile Gln His Ser Ser Ser Thr Leu Leu Gln Leu
          435          440          445
Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp
          450          455          460
Leu Glu Cys Val Glu Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val
465          470          475          480
Gln Ser Phe Thr Gly Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr
          485          490          495
Cys Leu Ser Ala Glu Leu Pro Leu Arg Met Arg Gly Ala Ala Ala Ser
          500          505          510
Ile Arg Gln Ile Leu Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr
          515          520          525
Asp Asn Gly Tyr Val Asn Val His Leu Lys Ala Ser Val Val Asp Ala
          530          535          540
Glu Cys Val Met Leu Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile
545          550          555          560
Asn Val Glu Asp Gln Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg
          565          570          575
Arg Ser Glu His Pro Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser
          580          585          590
Gln Arg Leu Ala Gln Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu
          595          600          605
Leu Gly Leu Gly Ser Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile
          610          615          620
Ala Met Gln Ala Glu Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val
625          630          635          640
Leu Ala Pro Val Arg Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser
          645          650          655
Arg Trp Gly Gly Arg Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu
          660          665          670
Ala Asp Ala Thr Ser Leu Leu Val Glu Val Leu Leu Leu Glu Gly Ala
          675          680          685
Pro Met Phe Glu Ala Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln
          690          695          700
Gly Asp Met Glu Pro Gln Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu
705          710          715          720
Asn Asn Leu Asp Gly Leu His Arg Ala Leu Gly Leu Ala His Gly Arg
          725          730          735
Leu Ala Asp Pro Ser Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn
          740          745          750

```

Leu Gly Leu Arg Val Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu
 755 760 765
 Ile Leu Arg Asp Gln Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu
 770 775 780
 Phe Asp Gly Arg Glu Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp
 785 790 795 800
 Val Val Leu Thr Asp Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu
 805 810 815
 Thr Ala Glu Leu Arg Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala
 820 825 830
 Thr Ala Asn Ala Met Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly
 835 840 845
 Met Asn Asp Cys Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn
 850 855 860
 Cys Leu Ile Asn Ile Leu Lys Val Asp Arg
 865 870

<210> 24

<211> 870

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 24

Met Tyr Glu Ala Gln Leu Tyr Phe Glu Gln Arg Glu Ala Leu Leu Asn
 1 5 10 15
 His Leu Ser Gly Asn Val Val Pro Leu Ala Ala Gly Arg Ala Leu Val
 20 25 30
 Asn Glu Ala Pro Asn Asn Val Ser Ile Leu Pro Leu Ser Asp Gly Gly
 35 40 45
 Arg Gly Leu Leu Leu Thr Ala Arg Thr Leu Gly Asp Leu Arg Glu Lys
 50 55 60
 Arg Leu Ala Leu Met Tyr Leu Val Asp Thr Asp Lys Gly Pro Leu Val
 65 70 75 80
 Tyr Arg Leu Thr Ala Asp Gly Arg Pro Ser Ala Ala Ile Ser Ser Thr
 85 90 95
 Ile Thr Lys Glu Val Tyr Arg Ala Leu Leu Ala Thr Pro Ser Ala Pro
 100 105 110
 Val His Trp Val Thr Asp Gly Gly Thr Pro Gln Arg Leu Tyr Leu Phe
 115 120 125
 Glu Ser Leu Gly Asp Glu Pro Gly Glu Gly Trp Leu Gly Leu Glu Ile
 130 135 140
 Leu Gly Glu Asp Leu Asp Ser Met Leu Arg Arg Asn Asp Ala Gly Asn
 145 150 155 160
 Tyr Met Leu Leu Asp Gln His Gly Gln Val Leu Ala Thr Asp Ala
 165 170 175
 Glu Ala Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp Gly
 180 185 190
 Phe Gly Phe Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu Phe
 195 200 205
 Gln His Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly Ile
 210 215 220
 Gly Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser Ala
 225 230 235 240
 Leu Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser Ile
 245 250 255
 Glu Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu Lys
 260 265 270

Glu Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val Ala
 275 280 285
 Leu Cys Val Leu Arg Arg Ala Asp Ala Val Val Leu Glu Asn Pro
 290 295 300
 Gln Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp Ala
 305 310 315 320
 Pro Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser Gly
 325 330 335
 Glu Glu Leu Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr Thr
 340 345 350
 Pro Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser Glu
 355 360 365
 Ile Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys Ser
 370 375 380
 Leu Ala Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr Met
 385 390 395 400
 Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu
 405 410 415
 Leu Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys
 420 425 430
 Ala Ile Gln His Ser Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp Val
 435 440 445
 Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys Val
 450 455 460
 Glu Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe Thr
 465 470 475 480
 Gly Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser Ala
 485 490 495
 Glu Leu Pro Leu Arg Met Arg Gly Ala Ala Ser Ile Arg Gln Ile
 500 505 510
 Leu Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly Tyr
 515 520 525
 Val Asn Val His Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val Met
 530 535 540
 Leu Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu Asp
 545 550 555 560
 Gln Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu His
 565 570 575
 Pro Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala
 580 585 590
 Gln Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu Gly
 595 600 605
 Ser Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln Ala
 610 615 620
 Glu Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro Val
 625 630 635 640
 Arg Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly
 645 650 655
 Arg Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr
 660 665 670
 Ser Leu Leu Val Glu Val Leu Leu Glu Gly Ala Pro Met Phe Glu
 675 680 685
 Ala Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met Glu
 690 695 700
 Pro Gln Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu Asn Asn Leu Asp
 705 710 715 720
 Gly Leu His Arg Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp Pro
 725 730 735

```

Ser Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu Arg
              740              745              750
Val Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg Asp
              755              760              765
Gln Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly Arg
              770              775              780
Glu Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp Val Val Leu Thr
785              790              795              800
Asp Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu Leu
              805              810              815
Arg Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn Ala
              820              825              830
Met Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp Cys
              835              840              845
Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile Asn
850              855              860
Ile Leu Lys Val Asp Arg
865              870

```

<210> 25

<211> 802

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 25

```

Met Tyr Leu Val Asp Thr Asp Lys Gly Pro Leu Val Tyr Arg Leu Thr
1              5              10              15
Ala Asp Gly Arg Pro Ser Ala Ala Ile Ser Ser Thr Ile Thr Lys Glu
              20              25              30
Val Tyr Arg Ala Leu Leu Ala Thr Pro Ser Ala Pro Val His Trp Val
              35              40              45
Thr Asp Gly Gly Thr Pro Gln Arg Leu Tyr Leu Phe Glu Ser Leu Gly
50              55              60
Asp Glu Pro Gly Glu Gly Trp Leu Gly Leu Glu Ile Leu Gly Glu Asp
65              70              75              80
Leu Asp Ser Met Leu Arg Arg Asn Asp Ala Gly Asn Tyr Met Leu Leu
              85              90              95
Asp Gln His Gly Gln Val Val Leu Ala Thr Asp Ala Glu Ala Leu Gly
              100              105              110
Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp Gly Phe Gly Phe Ile
              115              120              125
Gly Ala Gly Pro Leu Pro Gln His Met Val Leu Phe Gln His Val Gly
130              135              140
Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly Ile Gly Arg Leu Leu
145              150              155              160
Leu Ala Leu Trp Leu Pro Leu Leu Leu Ala Ser Ala Leu Ala Leu Ala
              165              170              175
Val Gly Ile Leu Leu His Trp Leu Val Arg Ser Ile Glu Arg Arg Leu
180              185              190
Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu Lys Glu Ser Glu Ala
195              200              205
Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val Ala Leu Cys Val Leu
210              215              220
Arg Arg Ala Asp Ala Val Val Leu Glu Asn Pro Gln Ala Arg Gln
225              230              235              240
Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp Ala Pro Arg Trp Ile
              245              250              255

```

```

Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser Gly Glu Glu Leu Glu
                260                265                270
Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr Thr Pro Thr Arg Tyr
                275                280                285
Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser Glu Ile Ser Ala Arg
                290                295                300
Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu Ala Asp Ala
305                310                315                320
Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr Met Ser His Glu Ile
                325                330                335
Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu Leu Gly Arg
                340                345                350
Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala Ile Gln His
                355                360                365
Ser Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp Val Leu Asp Val Ser
                370                375                380
Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys Val Glu Phe Ser Pro
385                390                395                400
Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe Thr Gly Ala Ala Gln
                405                410                415
Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser Ala Glu Leu Pro Leu
                420                425                430
Arg Met Arg Gly Ala Ala Ala Ser Ile Arg Gln Ile Leu Asn Asn Leu
                435                440                445
Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly Tyr Val Asn Val His
                450                455                460
Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val Met Leu Thr Trp Gln
465                470                475                480
Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu Asp Gln Pro Arg Leu
                485                490                495
Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu His Pro Val Ala Gly
                500                505                510
Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln Leu Met Asn
                515                520                525
Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu Gly Ser Ser Phe Ser
                530                535                540
Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln Ala Glu Pro Gln Asp
545                550                555                560
Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro Val Arg Asp Leu Thr
                565                570                575
Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg Ala Met Val
                580                585                590
Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser Leu Leu Val
                595                600                605
Glu Val Leu Leu Leu Glu Gly Ala Pro Met Phe Glu Ala Trp Pro Gly
                610                615                620
Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met Glu Pro Gln Ala Gln
625                630                635                640
Gly Arg Asp Trp Leu Leu Gly Leu Asn Asn Leu Asp Gly Leu His Arg
                645                650                655
Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp Pro Ser Thr Pro Pro
                660                665                670
Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu Arg Val Leu Val Val
                675                680                685
Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg Asp Gln Met Glu Ala
690                695                700                710
Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly Arg Glu Ala Leu Leu
705                710                715                720

```

His Cys Gln Thr Ala Cys Phe Asp Val Val Leu Thr Asp Ile Asn Met
 725 730 735
 Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu Leu Arg Arg Gln Gly
 740 745 750
 Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn Ala Met Arg Glu Glu
 755 760 765
 Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp Cys Leu Val Lys Pro
 770 775 780
 Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile Asn Ile Leu Lys Val
 785 790 795 800
 Asp Arg

<210> 26

<211> 719

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 26

Met Leu Arg Arg Asn Asp Ala Gly Asn Tyr Met Leu Leu Asp Gln His
 1 5 10 15
 Gly Gln Val Val Leu Ala Thr Asp Ala Glu Ala Leu Gly Ser Gly Ala
 20 25 30
 Ser Arg Thr Leu Leu Arg Gly Asp Gly Phe Gly Phe Ile Gly Ala Gly
 35 40 45
 Pro Leu Pro Gln His Met Val Leu Phe Gln His Val Gly Ser Ser Ser
 50 55 60
 Trp Asp Leu Ile Tyr His Ile Gly Ile Gly Arg Leu Leu Leu Ala Leu
 65 70 75 80
 Trp Leu Pro Leu Leu Leu Ala Ser Ala Leu Ala Leu Ala Val Gly Ile
 85 90 95
 Leu Leu His Trp Leu Val Arg Ser Ile Glu Arg Arg Leu Ile Glu Pro
 100 105 110
 Ala Lys Arg Arg Leu Glu Ala Leu Lys Glu Ser Glu Ala Phe Ser Arg
 115 120 125
 Ala Val Ile Gln Ala Ala Pro Val Ala Leu Cys Val Leu Arg Arg Ala
 130 135 140
 Asp Ala Ala Val Val Leu Glu Asn Pro Gln Ala Arg Gln Trp Leu Gly
 145 150 155 160
 Asp Ser Glu Ala Ile Ala His Asp Ala Pro Arg Trp Ile Ser Gln Ala
 165 170 175
 Phe Ala Gly Gly Val Lys Cys Ser Gly Glu Glu Leu Glu Thr Glu Ala
 180 185 190
 Gly Leu His Leu His Leu Asn Tyr Thr Pro Thr Arg Tyr Asn Gly Glu
 195 200 205
 Asp Val Leu Phe Cys Ala Phe Ser Glu Ile Ser Ala Arg Lys Arg Met
 210 215 220
 Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu Ala Asp Ala Ala Asn Glu
 225 230 235 240
 Ala Lys Thr Leu Phe Leu Ala Thr Met Ser His Glu Ile Arg Thr Pro
 245 250 255
 Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu Leu Gly Arg Thr Glu Leu
 260 265 270
 Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala Ile Gln His Ser Ser Ser
 275 280 285
 Thr Leu Leu Gln Leu Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu
 290 295 300

```

Ala Gly Gln Leu Asp Leu Glu Cys Val Glu Phe Ser Pro Leu Glu Leu
305          310          315          320
Thr Glu Glu Val Val Gln Ser Phe Thr Gly Ala Ala Gln Ala Lys Gly
          325          330          335
Leu Gln Leu Tyr Thr Cys Leu Ser Ala Glu Leu Pro Leu Arg Met Arg
          340          345          350
Gly Ala Ala Ala Ser Ile Arg Gln Ile Leu Asn Asn Leu Leu Ser Asn
          355          360          365
Ala Val Lys Phe Thr Asp Asn Gly Tyr Val Asn Val His Leu Lys Ala
          370          375          380
Ser Val Val Asp Ala Glu Cys Val Met Leu Thr Trp Gln Val Asn Asp
385          390          395          400
Thr Gly Met Gly Ile Asn Val Glu Asp Gln Pro Arg Leu Phe Glu Pro
          405          410          415
Phe Tyr Gln Ile Arg Arg Ser Glu His Pro Val Ala Gly Thr Gly Leu
          420          425          430
Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln Leu Met Asn Gly Ser Leu
          435          440          445
Lys Leu Val Ser Glu Leu Gly Leu Gly Ser Ser Phe Ser Leu Arg Leu
          450          455          460
Pro Leu Glu Arg Ile Ala Met Gln Ala Glu Pro Gln Asp Leu Ala Gly
465          470          475          480
Cys Ala Val Gln Val Leu Ala Pro Val Arg Asp Leu Thr Glu Cys Leu
          485          490          495
Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg Ala Met Val Ala Thr Pro
          500          505          510
Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser Leu Leu Val Glu Val Leu
          515          520          525
Leu Leu Glu Gly Ala Pro Met Phe Glu Ala Trp Pro Gly Cys Arg Val
          530          535          540
Glu Leu Ser Pro Gln Gly Asp Met Glu Pro Gln Ala Gln Gly Arg Asp
545          550          555          560
Trp Leu Leu Gly Leu Asn Asn Leu Asp Gly Leu His Arg Ala Leu Gly
          565          570          575
Leu Ala His Gly Arg Leu Ala Asp Pro Ser Thr Pro Pro Ile Arg Leu
          580          585          590
Ala Pro Leu Arg Asn Leu Gly Leu Arg Val Leu Val Val Glu Asp Asn
          595          600          605
Ala Ile Asn Gln Leu Ile Leu Arg Asp Gln Met Glu Ala Leu Gly Cys
          610          615          620
Ser Val Glu Leu Leu Phe Asp Gly Arg Glu Ala Leu Leu His Cys Gln
625          630          635          640
Thr Ala Cys Phe Asp Val Val Leu Thr Asp Ile Asn Met Pro Asn Met
          645          650          655
Asn Gly Tyr Glu Leu Thr Ala Glu Leu Arg Arg Gln Gly Phe Arg Gln
          660          665          670
Pro Ile Ile Gly Ala Thr Ala Asn Ala Met Arg Glu Glu Arg Glu Arg
          675          680          685
Cys Met Ser Ala Gly Met Asn Asp Cys Leu Val Lys Pro Val Asp Leu
690          695          700
Asn Ala Leu Gln Asn Cys Leu Ile Asn Ile Leu Lys Val Asp Arg
705          710          715

```

<210> 27

<211> 709

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 27

```

Met Leu Leu Asp Gln His Gly Gln Val Val Leu Ala Thr Asp Ala Glu
 1          5          10          15
Ala Leu Gly Ser Gly Ala Ser Arg Thr Leu Leu Arg Gly Asp Gly Phe
          20          25          30
Gly Phe Ile Gly Ala Gly Pro Leu Pro Gln His Met Val Leu Phe Gln
          35          40          45
His Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr His Ile Gly Ile Gly
          50          55          60
Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu Ala Ser Ala Leu
65          70          75          80
Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu Val Arg Ser Ile Glu
          85          90          95
Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu Glu Ala Leu Lys Glu
          100          105          110
Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala Ala Pro Val Ala Leu
          115          120          125
Cys Val Leu Arg Arg Ala Asp Ala Ala Val Val Leu Glu Asn Pro Gln
          130          135          140
Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile Ala His Asp Ala Pro
145          150          155          160
Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val Lys Cys Ser Gly Glu
          165          170          175
Glu Leu Glu Thr Glu Ala Gly Leu His Leu His Leu Asn Tyr Thr Pro
          180          185          190
Thr Arg Tyr Asn Gly Glu Asp Val Leu Phe Cys Ala Phe Ser Glu Ile
          195          200          205
Ser Ala Arg Lys Arg Met Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu
210          215          220
Ala Asp Ala Ala Asn Glu Ala Lys Thr Leu Phe Leu Ala Thr Met Ser
225          230          235          240
His Glu Ile Arg Thr Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu
          245          250          255
Leu Gly Arg Thr Glu Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala
          260          265          270
Ile Gln His Ser Ser Ser Thr Leu Leu Gln Leu Ile Ser Asp Val Leu
          275          280          285
Asp Val Ser Lys Ile Glu Ala Gly Gln Leu Asp Leu Glu Cys Val Glu
          290          295          300
Phe Ser Pro Leu Glu Leu Thr Glu Glu Val Val Gln Ser Phe Thr Gly
305          310          315          320
Ala Ala Gln Ala Lys Gly Leu Gln Leu Tyr Thr Cys Leu Ser Ala Glu
          325          330          335
Leu Pro Leu Arg Met Arg Gly Ala Ala Ala Ser Ile Arg Gln Ile Leu
          340          345          350
Asn Asn Leu Leu Ser Asn Ala Val Lys Phe Thr Asp Asn Gly Tyr Val
          355          360          365
Asn Val His Leu Lys Ala Ser Val Val Asp Ala Glu Cys Val Met Leu
          370          375          380
Thr Trp Gln Val Asn Asp Thr Gly Met Gly Ile Asn Val Glu Asp Gln
385          390          395          400
Pro Arg Leu Phe Glu Pro Phe Tyr Gln Ile Arg Arg Ser Glu His Pro
          405          410          415
Val Ala Gly Thr Gly Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln
          420          425          430
Leu Met Asn Gly Ser Leu Lys Leu Val Ser Glu Leu Gly Leu Gly Ser
          435          440          445
Ser Phe Ser Leu Arg Leu Pro Leu Glu Arg Ile Ala Met Gln Ala Glu

```



```

      450                      455                      460
Pro Gln Asp Leu Ala Gly Cys Ala Val Gln Val Leu Ala Pro Val Arg
465                      470                      475                      480
Asp Leu Thr Glu Cys Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg
      485                      490
Ala Met Val Ala Thr Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser
      500                      505                      510
Leu Leu Val Glu Val Leu Leu Leu Glu Gly Ala Pro Met Phe Glu Ala
      515                      520                      525
Trp Pro Gly Cys Arg Val Glu Leu Ser Pro Gln Gly Asp Met Glu Pro
      530                      535                      540
Gln Ala Gln Gly Arg Asp Trp Leu Leu Gly Leu Asn Asn Leu Asp Gly
545                      550                      555                      560
Leu His Arg Ala Leu Gly Leu Ala His Gly Arg Leu Ala Asp Pro Ser
      565                      570                      575
Thr Pro Pro Ile Arg Leu Ala Pro Leu Arg Asn Leu Gly Leu Arg Val
      580                      585                      590
Leu Val Val Glu Asp Asn Ala Ile Asn Gln Leu Ile Leu Arg Asp Gln
      595                      600                      605
Met Glu Ala Leu Gly Cys Ser Val Glu Leu Leu Phe Asp Gly Arg Glu
      610                      615                      620
Ala Leu Leu His Cys Gln Thr Ala Cys Phe Asp Val Val Leu Thr Asp
625                      630                      635                      640
Ile Asn Met Pro Asn Met Asn Gly Tyr Glu Leu Thr Ala Glu Leu Arg
      645                      650                      655
Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala Thr Ala Asn Ala Met
      660                      665                      670
Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly Met Asn Asp Cys Leu
      675                      680                      685
Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn Cys Leu Ile Asn Ile
      690                      695                      700
Leu Lys Val Asp Arg
705

```

<210> 28

<211> 666

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 28

```

Met Val Leu Phe Gln His Val Gly Ser Ser Ser Trp Asp Leu Ile Tyr
1      5      10      15
His Ile Gly Ile Gly Arg Leu Leu Leu Ala Leu Trp Leu Pro Leu Leu
      20      25      30
Leu Ala Ser Ala Leu Ala Leu Ala Val Gly Ile Leu Leu His Trp Leu
      35      40      45
Val Arg Ser Ile Glu Arg Arg Leu Ile Glu Pro Ala Lys Arg Arg Leu
      50      55      60
Glu Ala Leu Lys Glu Ser Glu Ala Phe Ser Arg Ala Val Ile Gln Ala
65      70      75      80
Ala Pro Val Ala Leu Cys Val Leu Arg Arg Ala Asp Ala Val Val
      85      90      95
Leu Glu Asn Pro Gln Ala Arg Gln Trp Leu Gly Asp Ser Glu Ala Ile
      100      105      110
Ala His Asp Ala Pro Arg Trp Ile Ser Gln Ala Phe Ala Gly Gly Val
      115      120      125
Lys Cys Ser Gly Glu Glu Leu Glu Thr Glu Ala Gly Leu His Leu His

```

-36-

```

      595              600              605
Thr Ala Glu Leu Arg Arg Gln Gly Phe Arg Gln Pro Ile Ile Gly Ala
      610              615              620
Thr Ala Asn Ala Met Arg Glu Glu Arg Glu Arg Cys Met Ser Ala Gly
625              630              635
Met Asn Asp Cys Leu Val Lys Pro Val Asp Leu Asn Ala Leu Gln Asn
      645              650              655
Cys Leu Ile Asn Ile Leu Lys Val Asp Arg
      660              665

```

<210> 29
 <211> 496
 <212> PRT
 <213> Pseudomonas aeruginosa PA14

```

<400> 29
Met Glu Ala Glu Leu Ala Arg Ala Lys Ser Leu Ala Asp Ala Ala Asn
 1      5      10      15
Glu Ala Lys Thr Leu Phe Leu Ala Thr Met Ser His Glu Ile Arg Thr
      20      25      30
Pro Leu Tyr Gly Met Leu Gly Thr Leu Glu Leu Leu Gly Arg Thr Glu
      35      40      45
Leu Ser Arg Gln Gln Ala Gly Tyr Leu Lys Ala Ile Gln His Ser Ser
      50      55      60
Ser Thr Leu Leu Gln Leu Ile Ser Asp Val Leu Asp Val Ser Lys Ile
      65      70      75      80
Glu Ala Gly Gln Leu Asp Leu Glu Cys Val Glu Phe Ser Pro Leu Glu
      85      90      95
Leu Thr Glu Glu Val Val Gln Ser Phe Thr Gly Ala Ala Gln Ala Lys
      100      105      110
Gly Leu Gln Leu Tyr Thr Cys Leu Ser Ala Glu Leu Pro Leu Arg Met
      115      120      125
Arg Gly Ala Ala Ala Ser Ile Arg Gln Ile Leu Asn Asn Leu Leu Ser
      130      135      140
Asn Ala Val Lys Phe Thr Asp Asn Gly Tyr Val Asn Val His Leu Lys
      145      150      155      160
Ala Ser Val Val Asp Ala Glu Cys Val Met Leu Thr Trp Gln Val Asn
      165      170      175
Asp Thr Gly Met Gly Ile Asn Val Glu Asp Gln Pro Arg Leu Phe Glu
      180      185      190
Pro Phe Tyr Gln Ile Arg Arg Ser Glu His Pro Val Ala Gly Thr Gly
      195      200      205
Leu Gly Leu Ser Ile Ser Gln Arg Leu Ala Gln Leu Met Asn Gly Ser
      210      215      220
Leu Lys Leu Val Ser Glu Leu Gly Leu Gly Ser Ser Phe Ser Leu Arg
      225      230      235      240
Leu Pro Leu Glu Arg Ile Ala Met Gln Ala Glu Pro Gln Asp Leu Ala
      245      250      255
Gly Cys Ala Val Gln Val Leu Ala Pro Val Arg Asp Leu Thr Glu Cys
      260      265      270
Leu Cys Gly Trp Ile Ser Arg Trp Gly Gly Arg Ala Met Val Ala Thr
      275      280      285
Pro Arg Ser Leu Asp Glu Ala Asp Ala Thr Ser Leu Leu Val Glu Val
      290      295      300
Leu Leu Leu Glu Gly Ala Pro Met Phe Glu Ala Trp Pro Gly Cys Arg
      305      310      315      320
Val Glu Leu Ser Pro Gln Gly Asp Met Glu Pro Gln Ala Gln Gly Arg

```

```
<210> 30
<211> 1992
<212> DNA
<213> Pseudomonas aeruginosa PA14
```

- 38 -

```

gtcgcgcacc tgcacgagga tgaccggcag gcgccccgct cgggcgttcg gcgaaggctc 1800
ggaagcgacc ctggtcagggt gcaccacatt ggcacgttcc tgcacgaggga ctctcctgcc 1860
accctcgagg cgcgcgatgg aatggcaaaa atcgggcaca gaggatcgat tggcgctcgtc 1920
cgtaacgtca atttccaggc gtcaaaaaca agtatctaca ttcattatag agatactttc 1980
aatctagat ag                                     1992

```

<210> 31

<211> 1830

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 31

```

atgccatttc tactggctac caagcacgcg ttgagcgccg attcctcggg agatccgttt 60
tcgctcgggtg tattgctcgc caatttctac ggaagcttct ggagtgtttc cgcctatccc 120
gcgccacagt tactgatctt tgatctttcc ggcagcaccg gcctggcagt gccgtcgatt 180
ccctccacag cgcagcgtga caggttgagc ggaagctatc cgatgatagt cgagcgatt 240
ctggcgcgct tgcgcacccg gccggtgggg gaggacgctc agcgtgtcca ttggatacgc 300
gctgatcgct atcgcgactc ggcgctggag atgttgggag tgcgccgggt tgatctgccg 360
gaaacactct ggtggcacga cgaagcgttg gttgagcgcc cggcattcga ttcgtacagc 420
ctcaggcgaa tcaatgactt cgaacagttg gttgagcgcc cggcattcga ttcgtacagc 480
ctggtatcgc cggatggcga ggtattgctc ggcgcggccc ctgcgaccgg cctgagggat 540
ggcctgaacc tcacccgaca gggggtcgcc gttcaactgc gcagccagcc tgagaacggc 600
tggctcgagg totaccgaac cgactacggc aatttcttcc gccactcccg gtgctgggtg 660
gcaggctcgc tgcgtacccc ggcgctgctc ctggccgggt ggctcgggat gcgttggtac 720
accagcagcg tcgtcaaccc ggtgcacgag gcgcaccggc aactggtgga gagcgacacc 780
ttcagccgga cgtgatata gaccgcgccc gtggctctgg tgggtgctgac ccaggatgac 840
cagcaactgg tgacctgcaa ccacttggcc gccagtgagg tggggcggcc cacggagatc 900
cttgggctga cttccaactg gaagcttttc gatgcgctg ggcaggatcc aggagacatc 960
tgtatccagg tcggtgggag ctatttgag accgccttc cggcgaccgg ctatgcggc 1020
accgagggcg tactgtgct attcaacgac atcacggtcc actgcgaggc ggagaccgag 1080
ctgtccaatg cgaagcgagc agcggatgcc gccagccagg ccaagaccct gttcctggcc 1140
cgcatgagcc atgaaatccg tactccctg tacggtgtcc ttggcaccct ggagtgtgctc 1200
gacctgacca ccctgaacga gcggcaacgc gcctacctac gcaccatcca gagttcgtct 1260
gcgacgctca tgcaactgat tagcgatgt ctggatgtct cgaagatcga agcggggcag 1320
atggctctga ccctggccgc cttcaatccg ctggacctag tgcgggaagt gcttggcaac 1380
tttgccgcca gcgccatggc caaggacctg caggtagacc cgctcgatac tcttgcgctt 1440
gaggcgagg tcgcgcgatg cttcgaagaa agcgttctgt tcgaggttgc tgggtggctc 1500
gtcggccatt tcgaagaggg tgcgtcgcc gttgtcgaac aacgcctgca acgcctgttt 1560
cagctgcagc gccgccttgt cgcgcacctg cagcaggatg accggcaggc gccccgctcc 1620
ggcggttcggc gacggctcgg aagcgaccct ggtcagggtg accacattgg catcgttctg 1680
catcgggact ctctgcccac cctcgcgccc gcgcattgaa tggcaaaaat cgggcacaga 1740
ggatcgattg gcgtcgctcc taacgtcaat ttccaggcgt caaaaaacaag tatctacatt 1800
cattatagag atactttcaa atctagatag                                     1830

```

<210> 32

<211> 1608

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 32

```

atgatagtcg agcgcattct ggcgcgcttg cgcacccggc cgggtggggga ggacgctcag 60
cgtgtccatt ggatacgcgc tgatcgctat cgcgactcgg cgctggagat gttgggagtc 120
gcccggttg atctgccgga aacactctgg tggcacgac agccgaacca tctgatcatc 180
gctgcgagcc tgcttgatct caggcgaatc aatgacttcc aacagttggt tgagcgccc 240
gcattcgatt cgtacagcct ggtatcgccc gatggcgagg tattgctcgg cgcgccccc 300
gcgaccggcc tgagggatgg cctgaacctc accgcacagg gggtcgccgt tcaactgcgc 360
agccagcctg agaacggctg gctcgcggtc taccgaaccg actacggcaa tttctttcgc 420
cactcccggg ggctggtggc aggtctgctg ctgaccccg cgctgctcct ggccggttgg 480

```

```

ctcgggatgc gttggtacac cagcagcgtc gtcaaccggt tgcacgggc gcaccggcaa 540
ctgggtggaga gcgacacctt cagccggagc ctgatacaga ccgcgccggt ggctctggtg 600
gtgctgaccc aggatgacca gcaactggtg acctgcaacc acttgccgc ccagtggctg 660
ggcggggccca cggagatcct tgggctgact tccaactgga agcttttcga tgcgcgtggg 720
caggtaccag gagacatctg tatccaggtc ggtgggcgct atttgcagac cgccttcgcg 780
gcgacccgct atgcggcac cgaggcggtg ctgtgcgtat tcaacgacat cacgggtccac 840
tgcgaggcgg agaccgcgct gtccaatgcg aagcgagcag cggatgccgc cagccaggcc 900
aagaccctgt tcctggcccg catgagccat gaaatccgta ctccctgta cgggtgtcctt 960
ggcaccctgg agttgctcga cctgaccacc ctgaacgagc ggcaacgcgc ctacctacgc 1020
accatccaga gttcgtctgc gacgtcatg caactgatta gcgatgtgct ggatgtctcg 1080
aagatcgaag cggggcagat ggctctgacc ctggcgcct tcaatccgct ggacctagt 1140
cgggaagtgc ttggcaactt tgccgcagc gccatggcca aggacctgca ggtagacctg 1200
ctcgatactc ttgcgcttga ggccgaggtc gcgcattggct tcgaagaaag cgttctgttc 1260
gaggttgctg gtggctcggc cggccatttc gaagagggtg tcgtcggcgt tgtcgaacaa 1320
cgctgcaac gcctgtttca gctgcagcgc cgcttctgct cgcacctgca cgaggatgac 1380
cggcaggcgc cccgctccgg cgctcggcga cggctcgga gcgacctgg tcaggtgcac 1440
cacattggca tcgttctgca tcgggactct cctgccacc tcgcggccgc gcatggaatg 1500
gcaaaaatcg ggcacagagg atcgattggc gtcgtccgta acgtcaattt ccaggcgtca 1560
aaaaaagta tctacattca ttatagagat actttcaaat ctagatag 1608

```

<210> 33

<211> 1500

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 33

```

atgttgggag tgcgccgggt tgatctgccg gaaacactct ggtggcacga cgagccgaac 60
catctgatca tcgctgcgag cctgcttgat ctacggcgaa tcaatgactt cgaacagttg 120
gttgagcgcc cggcattcga ttctgtacagc ctggatatgc cggatggcga ggtattgtct 180
ggcgccggcc ctgcgaccgg cctgagggat ggctgaacc tcaccgaca gggggtcgcc 240
gttcaactgc gcagccagcc tgagaacggc tggctcgccg tctaccgaac cgactacggc 300
aatctcttcc gccactcccg gtggctgggt gcaggtctgc tgctgacccc ggcgctgctc 360
ctggccgggt ggctcgggat gcgttggtac accagcagcg tcgtcaaccc ggtgcatcgg 420
gcgcaccggc aactggtgga gagcgacacc ttacgccgga cgctgataca gaccgcgccg 480
gtggctctgg tgggtgctgac ccaggatgac cagcaactgg tgacctgcaa ccacttgccc 540
gcccagtggt tgggcggggc cacggagatc cttgggctga cttccaactg gaagcttttc 600
gatgcgcgtg ggcaggtacc aggagacatc tgtatccagg tcggtgggcg ctatttgcag 660
accgccttcg cggcgaccgc ctatgccggc accgaggcgg tactgtgctt attcaacgac 720
atcacggtcc actgcgaggc ggagaccgcg ctgtccaatg cgaagcgagc agcggatgcc 780
gccagccagg ccaagaccct gttcctggcc cgcatgagcc atgaaatccg tactcccctg 840
tacggtgtcc ttggcaccct ggagtgtctc gacctgacca ccctgaacga gcggcaacgc 900
gcctacctac gcacctacca gagttcgtct gcgacgtca tgcaactgat tagcgatgtg 960
ctggatgtct cgaagatcga agcggggcag atggctctga ccctggccgc cttcaatccg 1020
ctggacctag tgccgggaagt gcttggcaac ttgcccga cgcgcattgg caaggacctg 1080
caggtagacc cgctcgatac tcttgcgctt gaggcgcagg tcgcgcatgg cttcgaagaa 1140
agcgttctgt tcgaggttgc tgggtggctc gtcggccatt tcgaagaggg tgtcgtcggc 1200
gttgtcgaac aacgcctgca acgcctgttt cagctgcagc gccgccttgt cgcgcacctg 1260
cacgaggatg accggcaggc gccccgctcc ggcgttcggc gacggctcgg aagcgacctt 1320
ggtcaggtgc accacattgg catcgttctg catcgggact ctctgccac cctcggcgcc 1380
gcgcatggaa tggcaaaaat cgggcacaga ggatcgattg gcgtcgtccg taacgtcaat 1440
ttccaggcgt caaaaaacaag tatctacatt cattatagag atactttcaa atctagatag 1500

```

<210> 34

<211> 1122

<212> DNA

<213> *Pseudomonas aeruginosa* PA14

<400> 34

```

atgcgttggg acaccagcag cgtcgtcaac ccggtgcatc gggcgccaccg gcaactgggtg 60
gagagcgaca ccttcagccg gacgctgata cagaccgcgc cgggtggctct ggtgggtgctg 120
acccaggatg accagcaact ggtgacctgc aaccacttgg ccgcccagtg gctgggagggg 180
cccacggaga tccttggggt gacttccaac tggaagcttt tcgatgcgcg tgggcaggta 240
ccaggagaca tctgtatcca ggtcgggtgg cgctatttgc agaccgcctt cgcggcgacc 300
cgctatgccg gcaccgaggg ggtactgtgc gtattcaacg acatcacggg cactgcgag 360
gaggagaccg cgctgtccaa tgcgaagcga gcagcggatg ccgccagcca ggccaagacc 420
ctgttcctgg ccgcgatgag ccatgaaatc cgtactcccc tgtacgggtg ccttggcacc 480
ctggagtgtc tcgacctgac caccctgaac gagcgggcaac gcgcctacct acgcaccatc 540
cagagttcgt ctgcgacgct catgcaactg attagcgatg tgctggatgt ctggaagatc 600
gaagcgggtc agatgggtct gaccctggcc gccttcaatc cgctggacct agtgcgggaa 660
gtgcttggca actttgccgc cagcgccatg gccaaaggacc tgcaggtaga cccgctcgat 720
actcttgcgc ttgagggcgca ggtcgcgat ggcttcgaag aaagcggtct gttcgaggtt 780
gctgggtggc cggtcggcca tttcgaagag ggtgtcgtcg gcgttgtcga acaacgcctg 840
caacgcctgt ttcagctgca gcgccgcctt gtcgcgcacc tgcacgagga tgaccggcag 900
gcgccccgct ccggcggttc gcgacggctc ggaagcgacc ctggtcaggt gcaccacatt 960
ggcatcgctt tgcacggga ctctcctgcc accctcgcgg ccgcgcgatg aatggcaaaa 1020
atcgggcaca gaggatcgat tggcgtcgtc cgtaacgtca atttccaggc gtcaaaaaca 1080
agtatctaca ttcattatag agatactttc aaatctagat ag 1122

```

<210> 35

<211> 663

<212> PRT

<213> *Pseudomonas aeruginosa* PA14

<400> 35

```

Met Asp Val Ile Arg Glu His Glu Val Phe Leu Gly Arg Ile Ala Arg
1          5          10          15
Lys Ser Asp Lys Thr Thr Gln Lys Tyr Asp Tyr Asp Val Val Pro Leu
20          25          30
Gln Arg His Leu Leu Ala Lys Glu Asn Gly Leu Ala Val Tyr Glu Gly
35          40          45
Arg Glu Phe Ser Phe Ala Met Pro Phe Leu Leu Ala Thr Lys His Ala
50          55          60
Leu Ser Ala Asp Ser Ser Gly Asp Pro Phe Ser Leu Gly Val Leu Leu
65          70          75          80
Ala Asn Phe Tyr Gly Ser Phe Trp Ser Val Ser Ala Tyr Pro Ala Pro
85          90          95
Gln Leu Leu Ile Phe Asp Leu Ser Gly Ser Thr Arg Leu Ala Val Pro
100          105          110
Ser Ile Pro Ser Thr Ala Gln Arg Asp Arg Leu Ser Gly Ser Tyr Pro
115          120          125
Met Ile Val Glu Arg Ile Leu Ala Arg Leu Arg Thr Arg Pro Val Gly
130          135          140
Glu Asp Ala Gln Arg Val His Trp Ile Arg Ala Asp Arg Tyr Arg Asp
145          150          155          160
Ser Ala Leu Glu Met Leu Gly Val Ala Arg Val Asp Leu Pro Glu Thr
165          170          175
Leu Trp Trp His Asp Glu Pro Asn His Leu Ile Ile Ala Ala Ser Leu
180          185          190
Leu Asp Leu Arg Arg Ile Asn Asp Phe Glu Gln Leu Val Glu Arg Pro
195          200          205
Ala Phe Asp Ser Tyr Ser Leu Val Ser Pro Asp Gly Glu Val Leu Leu
210          215          220
Gly Ala Ala Pro Ala Thr Gly Leu Arg Asp Gly Leu Asn Leu Thr Arg
225          230          235          240
Gln Gly Val Ala Val Gln Leu Arg Ser Gln Pro Glu Asn Gly Trp Leu

```

```

                245                250                255
Ala Val Tyr Arg Thr Asp Tyr Gly Asn Phe Phe Arg His Ser Arg Trp
                260                265                270
Leu Val Ala Gly Leu Leu Leu Thr Pro Ala Leu Leu Leu Ala Gly Trp
                275                280                285
Leu Gly Met Arg Trp Tyr Thr Ser Ser Val Val Asn Pro Val His Arg
                290                295                300
Ala His Arg Gln Leu Val Glu Ser Asp Thr Phe Ser Arg Thr Leu Ile
305                310                315                320
Gln Thr Ala Pro Val Ala Leu Val Val Leu Thr Gln Asp Asp Gln Gln
                325                330                335
Leu Val Thr Cys Asn His Leu Ala Ala Gln Trp Leu Gly Gly Pro Thr
                340                345                350
Glu Ile Leu Gly Leu Thr Ser Asn Trp Lys Leu Phe Asp Ala Arg Gly
                355                360                365
Gln Val Pro Gly Asp Ile Cys Ile Gln Val Gly Gly Arg Tyr Leu Gln
                370                375                380
Thr Ala Phe Ala Ala Thr Arg Tyr Ala Gly Thr Glu Ala Val Leu Cys
385                390                395                400
Val Phe Asn Asp Ile Thr Val His Cys Glu Ala Glu Thr Ala Leu Ser
                405                410                415
Asn Ala Lys Arg Ala Ala Asp Ala Ala Ser Gln Ala Lys Thr Leu Phe
                420                425                430
Leu Ala Arg Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Val Leu
                435                440                445
Gly Thr Leu Glu Leu Leu Asp Leu Thr Thr Leu Asn Glu Arg Gln Arg
                450                455                460
Ala Tyr Leu Arg Thr Ile Gln Ser Ser Ser Ala Thr Leu Met Gln Leu
465                470                475                480
Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Met Ala
                485                490                495
Leu Thr Leu Ala Ala Phe Asn Pro Leu Asp Leu Val Arg Glu Val Leu
                500                505                510
Gly Asn Phe Ala Ala Ser Ala Met Ala Lys Asp Leu Gln Val Asp Pro
                515                520                525
Leu Asp Thr Leu Ala Leu Glu Ala Gln Val Ala His Gly Phe Glu Glu
                530                535                540
Ser Val Leu Phe Glu Val Ala Gly Gly Ser Val Gly His Phe Glu Glu
545                550                555                560
Gly Val Val Gly Val Val Glu Gln Arg Leu Gln Arg Leu Phe Gln Leu
                565                570                575
Gln Arg Arg Leu Val Ala His Leu His Glu Asp Asp Arg Gln Ala Pro
                580                585                590
Arg Ser Gly Val Arg Arg Arg Leu Gly Ser Asp Pro Gly Gln Val His
                595                600                605
His Ile Gly Ile Val Leu His Arg Asp Ser Pro Ala Thr Leu Ala Ala
                610                615                620
Ala His Gly Met Ala Lys Ile Gly His Arg Gly Ser Ile Gly Val Val
625                630                635                640
Arg Asn Val Asn Phe Gln Ala Ser Lys Thr Ser Ile Tyr Ile His Tyr
                645                650                655
Arg Asp Thr Phe Lys Ser Arg
                660

```

<210> 36
 <211> 609
 <212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 36

```

Met Pro Phe Leu Leu Ala Thr Lys His Ala Leu Ser Ala Asp Ser Ser
 1          5          10          15
Gly Asp Pro Phe Ser Leu Gly Val Leu Leu Ala Asn Phe Tyr Gly Ser
          20          25          30
Phe Trp Ser Val Ser Ala Tyr Pro Ala Pro Gln Leu Leu Ile Phe Asp
          35          40          45
Leu Ser Gly Ser Thr Arg Leu Ala Val Pro Ser Ile Pro Ser Thr Ala
          50          55          60
Gln Arg Asp Arg Leu Ser Gly Ser Tyr Pro Met Ile Val Glu Arg Ile
          65          70          75          80
Leu Ala Arg Leu Arg Thr Arg Pro Val Gly Glu Asp Ala Gln Arg Val
          85          90          95
His Trp Ile Arg Ala Asp Arg Tyr Arg Asp Ser Ala Leu Glu Met Leu
          100          105          110
Gly Val Ala Arg Val Asp Leu Pro Glu Thr Leu Trp Trp His Asp Glu
          115          120          125
Pro Asn His Leu Ile Ile Ala Ala Ser Leu Leu Asp Leu Arg Arg Ile
          130          135          140
Asn Asp Phe Glu Gln Leu Val Glu Arg Pro Ala Phe Asp Ser Tyr Ser
          145          150          155          160
Leu Val Ser Pro Asp Gly Glu Val Leu Leu Gly Ala Ala Pro Ala Thr
          165          170          175
Gly Leu Arg Asp Gly Leu Asn Leu Thr Arg Gln Gly Val Ala Val Gln
          180          185          190
Leu Arg Ser Gln Pro Glu Asn Gly Trp Leu Ala Val Tyr Arg Thr Asp
          195          200          205
Tyr Gly Asn Phe Phe Arg His Ser Arg Trp Leu Val Ala Gly Leu Leu
          210          215          220
Leu Thr Pro Ala Leu Leu Leu Ala Gly Trp Leu Gly Met Arg Trp Tyr
          225          230          235          240
Thr Ser Ser Val Val Asn Pro Val His Arg Ala His Arg Gln Leu Val
          245          250          255
Glu Ser Asp Thr Phe Ser Arg Thr Leu Ile Gln Thr Ala Pro Val Ala
          260          265          270
Leu Val Val Leu Thr Gln Asp Asp Gln Gln Leu Val Thr Cys Asn His
          275          280          285
Leu Ala Ala Gln Trp Leu Gly Gly Pro Thr Glu Ile Leu Gly Leu Thr
          290          295          300
Ser Asn Trp Lys Leu Phe Asp Ala Arg Gly Gln Val Pro Gly Asp Ile
          305          310          315          320
Cys Ile Gln Val Gly Gly Arg Tyr Leu Gln Thr Ala Phe Ala Ala Thr
          325          330          335
Arg Tyr Ala Gly Thr Glu Ala Val Leu Cys Val Phe Asn Asp Ile Thr
          340          345          350
Val His Cys Glu Ala Glu Thr Ala Leu Ser Asn Ala Lys Arg Ala Ala
          355          360          365
Asp Ala Ala Ser Gln Ala Lys Thr Leu Phe Leu Ala Arg Met Ser His
          370          375          380
Glu Ile Arg Thr Pro Leu Tyr Gly Val Leu Gly Thr Leu Glu Leu Leu
          385          390          395          400
Asp Leu Thr Thr Leu Asn Glu Arg Gln Arg Ala Tyr Leu Arg Thr Ile
          405          410          415
Gln Ser Ser Ser Ala Thr Leu Met Gln Leu Ile Ser Asp Val Leu Asp
          420          425          430
Val Ser Lys Ile Glu Ala Gly Gln Met Ala Leu Thr Leu Ala Ala Phe

```

```

      435      440      445
Asn Pro Leu Asp Leu Val Arg Glu Val Leu Gly Asn Phe Ala Ala Ser
      450      455      460
Ala Met Ala Lys Asp Leu Gln Val Asp Pro Leu Asp Thr Leu Ala Leu
465      470      475      480
Glu Ala Gln Val Ala His Gly Phe Glu Glu Ser Val Leu Phe Glu Val
      485      490      495
Ala Gly Gly Ser Val Gly His Phe Glu Glu Gly Val Val Gly Val Val
      500      505      510
Glu Gln Arg Leu Gln Arg Leu Phe Gln Leu Gln Arg Arg Leu Val Ala
      515      520      525
His Leu His Glu Asp Asp Arg Gln Ala Pro Arg Ser Gly Val Arg Arg
      530      535      540
Arg Leu Gly Ser Asp Pro Gly Gln Val His His Ile Gly Ile Val Leu
545      550      555      560
His Arg Asp Ser Pro Ala Thr Leu Ala Ala Ala His Gly Met Ala Lys
      565      570      575
Ile Gly His Arg Gly Ser Ile Gly Val Val Arg Asn Val Asn Phe Gln
      580      585      590
Ala Ser Lys Thr Ser Ile Tyr Ile His Tyr Arg Asp Thr Phe Lys Ser
      595      600      605
Arg

```

```

<210> 37
<211> 535
<212> PRT
<213> Pseudomonas aeruginosa PA14

```

```

<400> 37
Met Ile Val Glu Arg Ile Leu Ala Arg Leu Arg Thr Arg Pro Val Gly
 1      5      10      15
Glu Asp Ala Gln Arg Val His Trp Ile Arg Ala Asp Arg Tyr Arg Asp
      20      25      30
Ser Ala Leu Glu Met Leu Gly Val Ala Arg Val Asp Leu Pro Glu Thr
      35      40      45
Leu Trp Trp His Asp Glu Pro Asn His Leu Ile Ile Ala Ala Ser Leu
      50      55      60
Leu Asp Leu Arg Arg Ile Asn Asp Phe Glu Gln Leu Val Glu Arg Pro
      65      70      75      80
Ala Phe Asp Ser Tyr Ser Leu Val Ser Pro Asp Gly Glu Val Leu Leu
      85      90      95
Gly Ala Ala Pro Ala Thr Gly Leu Arg Asp Gly Leu Asn Leu Thr Arg
      100      105      110
Gln Gly Val Ala Val Gln Leu Arg Ser Gln Pro Glu Asn Gly Trp Leu
      115      120      125
Ala Val Tyr Arg Thr Asp Tyr Gly Asn Phe Phe Arg His Ser Arg Trp
      130      135      140
Leu Val Ala Gly Leu Leu Leu Thr Pro Ala Leu Leu Leu Ala Gly Trp
      145      150      155      160
Leu Gly Met Arg Trp Tyr Thr Ser Ser Val Val Asn Pro Val His Arg
      165      170      175
Ala His Arg Gln Leu Val Glu Ser Asp Thr Phe Ser Arg Thr Leu Ile
      180      185      190
Gln Thr Ala Pro Val Ala Leu Val Leu Thr Gln Asp Asp Gln Gln
      195      200      205
Leu Val Thr Cys Asn His Leu Ala Ala Gln Trp Leu Gly Gly Pro Thr

```

```

      210      215      220
Glu Ile Leu Gly Leu Thr Ser Asn Trp Lys Leu Phe Asp Ala Arg Gly
225      230      235      240
Gln Val Pro Gly Asp Ile Cys Ile Gln Val Gly Gly Arg Tyr Leu Gln
      245      250      255
Thr Ala Phe Ala Ala Thr Arg Tyr Ala Gly Thr Glu Ala Val Leu Cys
      260      265      270
Val Phe Asn Asp Ile Thr Val His Cys Glu Ala Glu Thr Ala Leu Ser
      275      280      285
Asn Ala Lys Arg Ala Ala Asp Ala Ala Ser Gln Ala Lys Thr Leu Phe
      290      295      300
Leu Ala Arg Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Val Leu
305      310      315      320
Gly Thr Leu Glu Leu Leu Asp Leu Thr Thr Leu Asn Glu Arg Gln Arg
      325      330      335
Ala Tyr Leu Arg Thr Ile Gln Ser Ser Ser Ala Thr Leu Met Gln Leu
      340      345      350
Ile Ser Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Met Ala
      355      360      365
Leu Thr Leu Ala Ala Phe Asn Pro Leu Asp Leu Val Arg Glu Val Leu
      370      375      380
Gly Asn Phe Ala Ala Ser Ala Met Ala Lys Asp Leu Gln Val Asp Pro
385      390      395      400
Leu Asp Thr Leu Ala Leu Glu Ala Gln Val Ala His Gly Phe Glu Glu
      405      410      415
Ser Val Leu Phe Glu Val Ala Gly Gly Ser Val Gly His Phe Glu Glu
      420      425      430
Gly Val Val Gly Val Val Glu Gln Arg Leu Gln Arg Leu Phe Gln Leu
      435      440      445
Gln Arg Arg Leu Val Ala His Leu His Glu Asp Asp Arg Gln Ala Pro
      450      455      460
Arg Ser Gly Val Arg Arg Arg Leu Gly Ser Asp Pro Gly Gln Val His
465      470      475      480
His Ile Gly Ile Val Leu His Arg Asp Ser Pro Ala Thr Leu Ala Ala
      485      490      495
Ala His Gly Met Ala Lys Ile Gly His Arg Gly Ser Ile Gly Val Val
      500      505      510
Arg Asn Val Asn Phe Gln Ala Ser Lys Thr Ser Ile Tyr Ile His Tyr
      515      520      525
Arg Asp Thr Phe Lys Ser Arg
530      535

```

<210> 38

<211> 499

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 38

```

Met Leu Gly Val Ala Arg Val Asp Leu Pro Glu Thr Leu Trp Trp His
1      5      10      15
Asp Glu Pro Asn His Leu Ile Ile Ala Ser Leu Leu Asp Leu Arg
      20      25      30
Arg Ile Asn Asp Phe Glu Gln Leu Val Glu Arg Pro Ala Phe Asp Ser
      35      40      45
Tyr Ser Leu Val Ser Pro Asp Gly Glu Val Leu Leu Gly Ala Ala Pro
50      55      60
Ala Thr Gly Leu Arg Asp Gly Leu Asn Leu Thr Arg Gln Gly Val Ala

```

65					70					75				80
Val	Gln	Leu	Arg	Ser	Gln	Pro	Glu	Asn	Gly	Trp	Leu	Ala	Val	Tyr
				85					90					95
Thr	Asp	Tyr	Gly	Asn	Phe	Phe	Arg	His	Ser	Arg	Trp	Leu	Val	Ala
			100					105					110	Gly
Leu	Leu	Leu	Thr	Pro	Ala	Leu	Leu	Ala	Gly	Trp	Leu	Gly	Met	Arg
		115					120				125			
Trp	Tyr	Thr	Ser	Ser	Val	Val	Asn	Pro	Val	His	Arg	Ala	His	Arg
	130					135					140			Gln
Leu	Val	Glu	Ser	Asp	Thr	Phe	Ser	Arg	Thr	Leu	Ile	Gln	Thr	Ala
145					150					155				160
Val	Ala	Leu	Val	Val	Leu	Thr	Gln	Asp	Asp	Gln	Gln	Leu	Val	Thr
				165					170					175
Asn	His	Leu	Ala	Ala	Gln	Trp	Leu	Gly	Gly	Pro	Thr	Glu	Ile	Leu
		180						185					190	Gly
Leu	Thr	Ser	Asn	Trp	Lys	Leu	Phe	Asp	Ala	Arg	Gly	Gln	Val	Pro
	195						200					205		Gly
Asp	Ile	Cys	Ile	Gln	Val	Gly	Gly	Arg	Tyr	Leu	Gln	Thr	Ala	Phe
	210					215					220			Ala
Ala	Thr	Arg	Tyr	Ala	Gly	Thr	Glu	Ala	Val	Leu	Cys	Val	Phe	Asn
225					230					235				240
Ile	Thr	Val	His	Cys	Glu	Ala	Glu	Thr	Ala	Leu	Ser	Asn	Ala	Lys
				245					250					255
Ala	Ala	Asp	Ala	Ala	Ser	Gln	Ala	Lys	Thr	Leu	Phe	Leu	Ala	Arg
			260					265					270	Met
Ser	His	Glu	Ile	Arg	Thr	Pro	Leu	Tyr	Gly	Val	Leu	Gly	Thr	Leu
	275						280					285		Glu
Leu	Leu	Asp	Leu	Thr	Thr	Leu	Asn	Glu	Arg	Gln	Arg	Ala	Tyr	Leu
	290					295				300				Arg
Thr	Ile	Gln	Ser	Ser	Ser	Ala	Thr	Leu	Met	Gln	Leu	Ile	Ser	Asp
305					310					315				320
Leu	Asp	Val	Ser	Lys	Ile	Glu	Ala	Gly	Gln	Met	Ala	Leu	Thr	Leu
				325					330					335
Ala	Phe	Asn	Pro	Leu	Asp	Leu	Val	Arg	Glu	Val	Leu	Gly	Asn	Phe
			340					345					350	Ala
Ala	Ser	Ala	Met	Ala	Lys	Asp	Leu	Gln	Val	Asp	Pro	Leu	Asp	Thr
	355						360					365		Leu
Ala	Leu	Glu	Ala	Gln	Val	Ala	His	Gly	Phe	Glu	Glu	Ser	Val	Leu
	370					375				380				Phe
Glu	Val	Ala	Gly	Gly	Ser	Val	Gly	His	Phe	Glu	Glu	Gly	Val	Val
385					390					395				Gly
Val	Val	Glu	Gln	Arg	Leu	Gln	Arg	Leu	Phe	Gln	Leu	Gln	Arg	Arg
				405					410					415
Val	Ala	His	Leu	His	Glu	Asp	Asp	Arg	Gln	Ala	Pro	Arg	Ser	Gly
			420					425					430	Val
Arg	Arg	Arg	Leu	Gly	Ser	Asp	Pro	Gly	Gln	Val	His	His	Ile	Gly
	435						440					445		Ile
Val	Leu	His	Arg	Asp	Ser	Pro	Ala	Thr	Leu	Ala	Ala	Ala	His	Gly
	450					455				460				Met
Ala	Lys	Ile	Gly	His	Arg	Gly	Ser	Ile	Gly	Val	Val	Arg	Asn	Val
465					470					475				480
Phe	Gln	Ala	Ser	Lys	Thr	Ser	Ile	Tyr	Ile	His	Tyr	Arg	Asp	Thr
				485					490					495
Lys	Ser	Arg												

<210> 39

<211> 373

<212> PRT

<213> Pseudomonas aeruginosa PA14

<400> 39

```

Met Arg Trp Tyr Thr Ser Ser Val Val Asn Pro Val His Arg Ala His
 1           5           10           15
Arg Gln Leu Val Glu Ser Asp Thr Phe Ser Arg Thr Leu Ile Gln Thr
          20          25          30
Ala Pro Val Ala Leu Val Val Leu Thr Gln Asp Asp Gln Gln Leu Val
          35          40          45
Thr Cys Asn His Leu Ala Ala Gln Trp Leu Gly Gly Pro Thr Glu Ile
          50          55          60
Leu Gly Leu Thr Ser Asn Trp Lys Leu Phe Asp Ala Arg Gly Gln Val
65          70          75          80
Pro Gly Asp Ile Cys Ile Gln Val Gly Gly Arg Tyr Leu Gln Thr Ala
          85          90          95
Phe Ala Ala Thr Arg Tyr Ala Gly Thr Glu Ala Val Leu Cys Val Phe
          100         105         110
Asn Asp Ile Thr Val His Cys Glu Ala Glu Thr Ala Leu Ser Asn Ala
          115         120         125
Lys Arg Ala Ala Asp Ala Ala Ser Gln Ala Lys Thr Leu Phe Leu Ala
          130         135         140
Arg Met Ser His Glu Ile Arg Thr Pro Leu Tyr Gly Val Leu Gly Thr
145         150         155         160
Leu Glu Leu Leu Asp Leu Thr Thr Leu Asn Glu Arg Gln Arg Ala Tyr
          165         170         175
Leu Arg Thr Ile Gln Ser Ser Ser Ala Thr Leu Met Gln Leu Ile Ser
          180         185         190
Asp Val Leu Asp Val Ser Lys Ile Glu Ala Gly Gln Met Ala Leu Thr
          195         200         205
Leu Ala Ala Phe Asn Pro Leu Asp Leu Val Arg Glu Val Leu Gly Asn
          210         215         220
Phe Ala Ala Ser Ala Met Ala Lys Asp Leu Gln Val Asp Pro Leu Asp
225         230         235         240
Thr Leu Ala Leu Glu Ala Gln Val Ala His Gly Phe Glu Glu Ser Val
          245         250         255
Leu Phe Glu Val Ala Gly Gly Ser Val Gly His Phe Glu Glu Gly Val
          260         265         270
Val Gly Val Val Glu Gln Arg Leu Gln Arg Leu Phe Gln Leu Gln Arg
          275         280         285
Arg Leu Val Ala His Leu His Glu Asp Asp Arg Gln Ala Pro Arg Ser
          290         295         300
Gly Val Arg Arg Arg Leu Gly Ser Asp Pro Gly Gln Val His His Ile
305         310         315         320
Gly Ile Val Leu His Arg Asp Ser Pro Ala Thr Leu Ala Ala Ala His
          325         330         335
Gly Met Ala Lys Ile Gly His Arg Gly Ser Ile Gly Val Val Arg Asn
          340         345         350
Val Asn Phe Gln Ala Ser Lys Thr Ser Ile Tyr Ile His Tyr Arg Asp
          355         360         365
Thr Phe Lys Ser Arg
          370

```